نماذج مختارة
في الإحصاء الجغرافي

ا. د. امحمد عياد محمد مقيلي
استاذ جغرافية المناخ بقسم الجغرافيا
كلية الآداب-جامعة طرابلس

منشورات مركز البحوث والاستشارات - جامعة سرت
الطبعة الأولى 2025

نماذج مختارة

في الإحصاء الجغرافي

أ. د/ إمحمد عياد محمد مقيلي

قسم الجغرافيا/ كلية الآداب/ جامعة طرابلس

منشورات مركز البحوث والاستشارات بجامعة سرت

الطبعة الأولى 2025م

نماذج مختارة في الإحصاء الجغرافي

تأليف
أ. د. إمحمد عياد محمد مقيلي
قسم الجغرافيا/ كلية الآداب/ جامعة طرابلس

دار الكتب الوطنية ـ بنغازي
رقم الإيداع القانوني: **719 /2024** م
رقم الإيداع الدولي
ردمك 3-3519-1-9959-978 ISBN

تصميم الغلاف: **خالد جمعة امهلهل**

منشورات مركز البحوث والاستشارات بجامعة سرت
الطبعة الأولى: 2025م

بسم الله الرحمن الرحيم

﴿وَكُلَّ شَيْءٍ أَحْصَيْنَاهُ كِتَابًا﴾

صدق الله العظيم

الآية 29 سورة النبأ

الإهداء

إلى كل من أحب العلم وتفانى في تعلمه

وتعليمه أهدي هذا الكتاب

مقيلي

المحتويات

المحتويات

المحتويات

تقديم:

تهتم علوم الجغرافيا بفروعها الطبيعية والبشرية بالمشاكل المتجددة التي يواجهها الإنسان في حياته اليومية، ويحاول من يبحث في هذا المجال أن يجد لها حلولا تخفف من آثارها السلبية على حياته وعلى بيئته التي يعيش في إطارها. وتتعدد المشاكل فمنها الناتج عن ظروف طبيعية متطرفة كالزلازل والأعاصير المسببة للفيضانات والتبدل المناخي المتسبب في دورات الجفاف والحرائق والأوبئة الفتاكة ومنها ناتجا عن الزيادة السكانية والتوسع العمراني والمنافسة الدولية على موارد محدودة ومنها غير المتجدد كالوقود الأحفوري وبعض المعادن. وينشأ عن التوسع الزراعي والرعوي مشاكل مرتبطة بظواهر التصحر والانقراض الحيوي. وينشأ عن النمو السكاني مشاكل ارتفاع أسعار السلع الغذائية وقلة فرص العمل والجريمة والازدحام المروري والحوادث المميتة. وينشأ عن التنافس على الموارد القلاقل الداخلية والحروب بين الدول، وينشأ عن التوسع الصناعي غير المضبوط مشاكل التلوث البيئي الخطر في البر والبحر والجو والغذاء. وعندما تشح الموارد وتضيق الأرض بسكانها لا يجد الكثير من الشباب مناصا من الهجرة متوقعين فرص حياة أفضل، لذلك فالعالم منذ فترة يعاني من مشكلة الهجرة غير الشرعية ومن مشاكل الفقر وتفشي الأمراض الوبائية بين الفينة والأخرى نتيجة هجرة الأفراد والنقل السريع بالطائرات.

كل تلك المشاكل التي ذكرت، والكثير غيرها لم يذكر، لعدم اتساع المجال لذكرها تحتاج إلى دراسات علمية تستخدم النماذج الإحصائية في الاختبارات والتنبؤ

بجدوى الحلول المقترحة بناء على الفرضيات الموضوعة عن أسباب تلك المشاكل. لهذا جاءت صفحات هذا الكتاب لكي تعالج الأسلوب العلمي في إجراء البحوث الجغرافية والتي تبدأ بالشعور بالمشكلة وبوضع الفروض المنطقية لمسبباتها والتأكد من صحة تلك الفروض بإجراء الاختبارات على البيانات الموثقة والتأكد من صحتها أو رفضها بناء على مقاييس الاختبارات المحددة بمستوى ثقة عالية. لقد جاءت فصول هذا الكتاب متتابعة تبين للطالب الاختبارات الممكن أن تناسب مشكلته البحثية في تحليل البيانات لأجل قبول أو رفض الفرضيات التي يشعر بها كحل لمشكلة الدراسة.

هذا الكتاب لم يأتي نتيجة جهد فكرة آنية ونفذت في أيام وأشهر قليلة وإنما هي أفكار متراكمة خلال تدريس مقرر عقود من الزمن تزيد عن الأربعة رأيت أن أجمعها في كتاب منهجي واحد حتى لا يأكل عليها الدهر ويشرب وتضيع. في سنة 2017 تقاعدت عن التدريس الجامعي وطلب مني زملائي في قسم الجغرافيا بإلحاح العودة إلى التدريس والإشراف على طلاب الدراسات العليا في مرحلتي الماجستير والدكتوراه، وأخص بالذكر منهم: الدكتور ناجي الزناتي، والدكتور إبراهيم فايد، والدكتور عمر الهاشمي يوسف، والدكتورة سميره العياطي، والدكتوره كريمة الدويب، ومن جامعة الزاوية الدكتوره إسمهان عثمان، ومن جامعة سرت الدكتور حسين أبومدينة، الذي اقترح إمكانية نشر عدد من كتبي بجامعة سرت الموقرة . وأخيرا أراد الله أن أعود فقبلت في الفصل الدراسي خريف 2023- 2024 الرجوع فوجدت همة لم أكن أعهدها من النشاط وكان أكبر محفز لي مجموعة من طلبة الدكتوراه المتميزين المتشوقين للمعرفة فرأيت أنهم جديرون بأن يحملوا رسالة الاستمرارية في نقل المعرفة التي جمعتها طيلة العقود الأربعة فأخذت على عاتقي تنظيم المقررات التي كنت أدرسها في أربعة كتب هي: مناخ العالم، والمناخ التطبيقي، والإحصاء الجغرافي، ومشاكل البيئة أرجوا أن تجد طريقها للنشر ومن بينها هذا الكتاب.

تناول الفصل الأول من هذا الكتاب ماهية المشكلة البحثية وفرضياتها. وعالج الفصل الثاني الإمكانيات التي يوفرها التوزيع التكراري الاعتدالي في التنبؤ بخصوص مجتمع العينة المدروسة. أما الفصل الثالث فيتناول الاختبارات التي تجرى على المتوسطات المتعلقة بمجتمع واحد خلال زمنين مختلفين وعلى مجتمعين في نفس الزمن، والاختبارات على النسب بدلا من الأعداد في حالة عدم تساوي أعداد العينات. وعالج الفصل الخامس مسألة الاختبارات اللامعلمية كحل لمشكلة صغر حجم العينة وانتفاء اعتدالية التوزيع لمفرداتها. واختص الفصل السادس بتحليل التباين بين أكثر من مجتمعين إحصائيين وتطرق الفصل السابع إلى تحليل المسلسلة الزمنية وأسلوب الخمس سنوات المتدحرجة ونموذج باينوميال لكشف الدورات الطويلة في مسلسلة البيانات المدروسة. وزود كل فصل بتمرينات إضافية يقوم الطالب بتحليها في البيت لترسيخ ما فهمه وطبقة عمليا في الفصل. وأخيرا أرجو أن أكون قد وفقت في تقديم ما ذكرت وإن وفقت فهو منة من الله فهو ولي التوفيق.

أ. د. إمحمد عياد محمد مقيلي

طرابلس – ليبيا

20 سبتمبر 2024م.

الفصل الأول

المشكلة البحثية وفرضياتها

1- المشكلة البحثية:

تعتري الباحث في تخصصه أفكار بخصوص تصحيح وضع مقلق ربما يكون مسببا لخسائر أو تعطل في العمل، أو لا يسبب خسائر أو عرقلة وإنما يريد الباحث تطوير الأداء إلى وضع أفضل مريح. حياة الإنسان عبارة عن تفاعل مستمر بينه وبين وما حوله من بشر ومواد وموارد، وينتج عن هذا التفاعل الكثير من المشاكل التي تعرقل صفو حياته وتتسبب له الخسائر لذلك لزم حلها من أجل تحقيق الراحة والطمأنينة والتقدم. وفيما يلي بعض المشاكل : تدني خدمات الصرف الصحي أو غيابها عن كثير من المناطق الحضرية، التوتر الدولي والحروب، تغير المناخ، غلاء المعيشة، العزوف عن الزواج ، ارتفاع معدلات الطلاق، انتشار الجريمة، ظهور أمراض جديدة مستعصية عن العلاج، ارتفاع تكلفة العلاج ، تدني مخرجات التعليم، ازدحام المرور، وتدني المستوى التعليمي والأمني، مشاكل الجفاف والتصحر إلى غير ذلك من المواضيع الممكن البحث فيها.

2- الهدف من الدراسات العلمية

يعرف الهدف من الدراسات العلمية بكل بساطة على أنه السعي لإيجاد حلول للمشكلات التي تواجه الإنسان وتؤثر سلبًا على حياته. إنها مجهود مضني لا بد منه لأجل التخطيط السليم لمشاريع التنمية التي تحسن التقدير لأحجام المشروعات ولا تبالغ فيها أو تبخسها.

إن التعليم العالي معرفة منظمة وبحث مستمر عن الحقائق بغية الوصول إلى معرفة قابلة للاختبار بواسطة التجربة. الهدف الأسمى للعلوم هو تطوير السلوك والخدمات والأساليب والإنتاج وذلك من خلال تحويل الجهد العلمي والمهارات الإنسانية في النظم والبحوث إلى نواتج مادية لكي تسهم في تحقيق الرفاهية المادية والنفسية للإنسان. إن نتائج العلم هو المعرفة في الاستخدام والتطبيق مستندا على التقنية وهي التي تمثل الوسيلة في تصنيع التكنولوجيا وإدخالها في التنمية. (حميد،2007، ص12)

إن الغاية من البحث العلمي الجاد هو التعمق في المعرفة والبحث عن الحقيقة والسعي إلى تزويد المجتمع بالمعرفة والعلم، والمساهمة الإيجابية في تقديم الحلول لمشكلاته وإيجاد العلاج الناجح لها. إن ميادين البحث العلمي لم تعد قاصرة على العلوم التطبيقية وحدها، فهي امتدت لتشمل جميع المشكلات والظواهر التي يتعرض لها المجتمع، وأصبح أداة للوصول إلى الحقيقة في مختلف المجالات مما سينتج عنه تجديد ذلك العلم وتطوره، وبالتالي تطور ورقي المجتمع، وهذا لا يتأتى إلا من خلال تحقيق الشخصية الذاتية للباحث، بمعنى أننا لا نجتر ما قام به غيرنا بل نستفيد منه ونحاول أن نضيف إليه ونطوره وأن نبتكر أفكارا جديدة لم يتطرق إليها أحد من قبل ونختبرها لنتأكد من مصداقيتها على أرض الواقع. (شلبي، 1989، ص13)

وتتوقف قيمة العمل العلمي على عدة عوامل أهمها أن يكون هدف الباحث استجلاء الحقيقة وإظهارها، سواء اتفقت مع ميوله وتوجهاته، أو لم تتفق. ولتحقيق ذلك لا بد من موافاة الشروط التالية:

1- الاطلاع الواسع حول الموضوع المراد البحث فيه بالتنقيب في الكتب والبحوث المنشورة في الدوريات المتخصصة ورسائل الماجستير والدكتوراه وأعمال المؤتمرات العلمية.

2- أن يكون الباحث ملما بلغة أجنبية متداولة كالإنجليزية والفرنسية والإسبانية والألمانية حتى يتمكن من الاطلاع على أحدث البحوث المنشورة في المجال الذي يدرسه.

3- أن يكون الباحث متمكنا من أساليب البحث والتحليل العلمي.

4- أن يفهم فهما دقيقا آراء غيره وألا يأخذها كحقيقة مسلم بها بل يخضعها للتحليل والنقاش وأن يوضح نقاط الاختلاف ونقاط الاتفاق مع ما توصل إليه في دراسته مبينا السبب في ذلك الاختلاف إن وجد. فالاختلاف وارد ربما نتيجة لاستخدام بيانات جديدة أو وسائل مطورة في التحليل.

5- أن يضيف البحث شيئا جديدا وأن يبتعد عن اجترار ما كتبه غيره.

6- يجب أن تظهر شخصية الباحث في بحثه من خلال الكتابة باللغة العربية السليمة وأن يكون متمكنا من التعبير عن أفكاره بجمل واضحة مركزة لا غموض فيها.

7- على الباحث التحلي بالصبر والفطنة والدقة والحياد، وألا يستعجل النتائج من أجل التخرج والحصول على وظيفة.

3- الفرضية العلمية:

هي جملة STATEMENT منطقية يضعها الباحث كتفسير مبدئي لسبب المشكلة. إنه تخمين منطقي لسبب المشكلة ويحتمل أن يكون صحيحا أو خطأ ونتائج التحليل تثبت ذلك أو ترفضه. لنفرض أن مدرسة ما مشهورة بتدني مستوى خريجيها وعدم نجاحهم بالجامعة، وسئل مجموعة من المدرسين والموجهين عن سبب هذا التخلف، فإجاباتهم لن تخرج عن الفرضيات الآتية:

1- تدني مستوى المعلمين نتيجة لتغير المناهج أو نتيجة لتكليف بعضهم تدريس مواد خارج تخصصهم، أو نتيجة لعدم تأهيلهم من الأساس للمهنة.

2- التخلف الثقافي لأولياء الأمور يجعلهم لا يتابعون مسيرة أبنائهم في المدرسة ولا متابعة الواجبات المنزلية لأبنائهم ولا يلبون متطلباتهم من كتب ووسائل يحتاجونها.

3- تفشي البطالة في المنطقة يجعل الأولياء فقراء ليس لديهم القدرة المالية لتلبية مستلزمات أبنائهم.

4- الوضع الدراسي في المدرسة غير مريح نتيجة لارتفاع مستوى الضوضاء من طريق مزدحم بالسيارات أو من مصنع أو ورشة قريبة أو مطار. كما أن ندرة الوسائل المساعدة في العملية التعليمية يتسبب في عدم الفهم لذي التلاميذ.

5- صعوبة المواصلات تجعل البعض يصل متأخرا في الأيام العادية ويكثرون الغياب في الظروف الجوية القاسية.

أما أسباب حوادث المرور على الطرقات فهي متعددة ويمكن حصر أهمها في النقاط الآتية:

1- القيادة المتهورة بسرعة عالية يفقد السائق قدرة التحكم في المركبة التي يقودها.

2- ازدحام الطريق يوتر السائق عصبيا ويقلل انتباهه للسيارات التي تضايقه من جميع الاتجاهات.

3- عدم انتباه السائق نتيجة لانشغاله بأحد الأطفال المرافقين له أو بالهاتف النقال.

4- عدم الالتزام بقواعد المرور، كالاجتياز من على اليمين، أو الاقتراب غير الآمن من السيارات الأخرى أو الاجتياز في الأماكن الممنوع الاجتياز فيها.

6- القيادة تحت تأثير مخدر يفقد العقل قدرة التركيز.

7- القيادة بدون ترخيص.

8- ظروف جوية غير ملائمة مثل سقوط الأمطار أو الثلوج أو وجود الضباب، أو التعرض للعواصف الغبارية الحاجبة للرؤية.

9- تخلف البنية التحتية للطريق كغياب الإنارة ، وضيق الطريق ، وكثرة الانعطافات، والحفر.

10- عدم ملائمة السيارة للقيادة كضعف الكوابح، غياب الإضاءة، وتدهور حالة المحرك.

وهناك فرضيات أخرى يمكن سردها لتفسير أسباب وقوع الحوادث المرورية وعلى المحقق اختبار هذه الفرضيات لتحديد أيا منها صحيحة فيبني على أساسها تقريره الذي سيقدم للنيابة العامة والمحكمة التي ستدين المذنب وتوقع عليه العقوبة.

القرار العشوائي: هو أمر تنفيذي لم يسبقه دراسة وينتج عنه أحد الحالات التالية:

1- مبالغة التقدير Over estimation والتصميم لمشروع كبير جدا لا يستغل بالكامل ويتطلب إنشائه صرف مبالغ طائلة من الأموال، ويحتاج لجهد كبير ووقت طويل ليتم انجازه(إنه مضيع للوقت والجهد والمال).

2- ابخاس التقدير Under estimation والتصميم لمشروع صغير موفر للوقت والجهد والمال لكنه لا يفي بالغرض المقام لشأنه ولا يحل المشكلة.

3- حسن التقدير Good estimation التصميم لمشروع مثالي بالصدفة.

القرار العلمي: هو أمر تنفيذي مبني على دراسة مستوفية لجميع المؤثرات وثبت نجاحها بالاختبارات اللازمة مما يضمن نجاح المشروع ويحقق أكبر قدر من الفائدة ويجعل الخسائر في حدها الأدنى الذي يمكن تحمله.

4- مسار التفكير العلمي.

يمر البحث العلمي بعدد من الخطوات المتتابعة وهي:

1- تحديد المشكلة.

2- وضع الفرضيات المفسرة.

3- جمع البيانات المتعلقة بالمشكلة.

4- تحليل البيانات وإجراء الاختبارات للتحقق من الفرضيات.

5- الخروج بنتائج وتوصيات لحل المشكلة.

ويقود التخطيط المبني على أسلوب البحث العلمي إلى حسن التقدير للمشروعات مباشرة دون التعرض لمخاطر الصدفة التي تقود أحيانا لمبالغة التقدير وأحيانا أخرى لإبخاس التقدير التي يتصف بها التفكير العشوائي.

5- مصادر البيانات والمعلومات.

تتعدد مصادر البيانات والمعلومات التي يمكن أن يستفيد منها الباحث لإنجاز عمله ويمكن تصنيفها في النقاط التالية:

1- سجلات مؤسسات الدولة التي توثق الداخل والخارج من المعطيات التي تخصها فهناك على سبيل المثال سجلات المستشفيات والمصحات، والبريد، والموانئ والمطارات ونقاط الحدود البرية، الغرف التجارية، الأرصاد الجوية، وسجلات مراكز البحوث الزراعية والصناعية والدوائية والنفط، وسجلات البلديات ومراكز الشرطة والمحاكم وسجلات الوزارات كالتعليم والرياضة والإسكان والمواصلات والجوازات والهجرة إلى غير ذلك.

2- سجلات المصانع والشركات الصناعية والتجارية.

3- مطبوعات منظمات الأمم المتحدة مثل منظمة الأمومة والطفولة، والأغذية والزراعة، والأرصاد العالمية، والطيران الدولي، والسكان.

4- بيانات يجمعها الباحث لوحده عن طريق استبيانات استطلاع الرأي والمقابلات الشخصية من المتخصصين وأصحاب الشأن أو عن طريق القياس، أو أخذ عينات وتحليها في المختبرات، أو عن طريق عمل اختبارات موضوعية على عينات من مجتمع الدراسة.

6- المجتمع الإحصائي والعينة والمتغير:

تهدف الدراسات العلمية إلى معرفة خصائص الظاهرة قيد الدراسة من ناحية التوزيع والحركة عبر الزمن والعوامل المؤثرة فيها وكيف تؤثر بدورها في حياة الإنسان لأجل التوصل إلى تنبؤات مستقبلية سليمة تفيد في اتخاذ قرارات صحيحة تقلل من الآثار السلبية للظاهرة.

افرض إن مصنعا أنتج 10000 مصباحا كهربائيا، ويدعي بأن العمر التشغيلي لهذه المصابيح يصل إلى 2000 ساعة، فكيف يمكن لمركز البحوث الصناعية التحقق من صحة هذا الادعاء من خطأه. طبعا الطريقة المؤكدة للتحقق من ذلك هي تشغيل كل ما أنتجه المصنع (10000) مصباح لمدة 2000 ساعة، وفي هذه الحالة سوف يستهلك كل ما أنتجه المصنع في اختبارات ولن يخرج للسوق شيئا، وهذا بالطبع إسراف لا مبرر له. أما الطريقة المعقولة والأسلم

فتتم بأخذ عينة عشوائية من المصابيح لنقل (50) قطعة على سبيل المثال، وإجراء الدراسة عليها.

المجتمع الطلابي في الجامعة يعني جميع أفراد الطلبة والطالبات، أما العينة الطلابية فهي مجموعة من الطلبة والطالبات مختارة من المجتمع الطلابي ولتكن 500 طالب وطالبة على سبيل المثال. وعندما نختار عينة وندرسها فنحن لا نهدف إلى معرفة خصائص العينة بحد ذاتها، بل نهدف إلى معرفة خصائص المجتمع المأخوذة منه. ولهذا تعتمد درجة التوفيق في معرفة خصائص المجتمع في طريقة اختيارنا للعينة، فلكي تكون العينة ممثله بدرجة كبيرة للمجتمع و غير منحازة لفئة دون أخرى، يجب أن يكون اختيارنا لها عشوائيا، ويجب أن يكون عدد أفرادها كبيرا بقدر الإمكان، لأنه مع زيادة المفردات (حجم العينة) تزداد المعلومات الممكن التوصل إليها عن المجتمع.

إن بيانات قليلة تعطي فكرة خاطئة عن وجود تغير بالرغم من عدم وجوده على المدى الطويل، فهذا التغير لذى البيانات القصيرة قد يظهر إلا ذبذبة من الذبذبات العشوائية على المدى الطويل. فإذا اكتفينا بقرائتين للحرارة (على سبيل المثال) ليوم 4 ابريل فنجدها 30 و 36 لعامي 1993 و 1994 على التوالي، فهل هذا يعني أن درجة الحرارة متجهة نحو الزيادة أو إن هناك اتجاه مناخي نحو ارتفاع الحرارة في العالم ؟. طبعا لا، فالاتجاه ألمناخي لا نتعرف عليه إلا من خلال دراسة سجل الحرارة لمدة طويلة.

التغير الحاصل في درجة حرارة يوم 4 ابريل بين السنتين ربما يكون ناتجا عن استبدال الترمومتر القديم بآخر جديد، أو ناتجا عن تغير في الظروف المحيطة بكشك الأرصاد كتركيب مكيف يضخ الهواء الساخن باتجاه الكشك أو قد يكون ناتجا عن اختلاف في خصائص الكتل الهوائية المرتحلة. فكل هذه المتغيرات تتسبب في حصول تباين في قياسات الحرارة لنفس الموقع بين السنوات ولا يمكن إرجاعها إلى تغيرات مناخية، ومن ثم فان معلوماتنا تزداد بالتأكيد مع زيادة حجم العينة لذلك فمن الأفضل دراسة عينات متكونة من أكبر عدد ممكن من المفردات بدلا من الاكتفاء بعينات صغيرة وتوقع نتائج مبهرة منها.

إن العشوائية في اختيار العينة ضرورية جدا، افرض (على سبيل المثال) إن عدد طلبة الجامعة متكون من 10000 نسمة منهم 5000 طالب و5000 طالبة، فإذا أخذنا عينة متكونة من 100 نسمة، فسوف يكون اختيارنا لها موفق إذا كانت متكونة من 50 طالب و50 طالبة أو ما يقرب من ذلك. أما إذا انحزنا إلى أحد الفريقين فان العينة سوف تبتعد عن حقيقة المجتمع المتكون أساسا من 50 ٪ ذكور و 50 ٪ إناث.

مما تقدم يتضح أن المجتمع الإحصائي population يعني جميع مفردات أو مشاهدات الظاهرة المدروسة مثل سكان ليبيا، أو سكان طرابلس، أو السيارات الملاكي، أو صادرات النفط أو حوادث المرور، أو مسلسلات المطر إلى غير ذلك. أما ألعينة الإحصائية statistic sample فهي عدد مختار من المشاهدات لغرض إجراء الدراسة عليها والتوصل إلى معلومات يفترض أن تخص المجتمع ككل. ولأجل الحصول على عينة ممثلة للمجتمع يجب أن يكون اختيارنا لها عشوائيا random sample أي غير منحازين لطرف دون آخر كما يجب أن تكون كبيرة الحجم بقدر الإمكان. لأنه كلما زاد حجم العينة كلما زادت دقة النتائج المتحصل عليها بخصوص المجتمع. فلا يكفي دراسة عدد قليل من مفردات المجتمع ونتوقع نتائج مبهرة، فالعينة الصغيرة ربما تقود إلى نتائج مغالطة للواقع لأنها لن تحتوي على كل خصائص المجتمع المأخوذة منه. كما أن عشوائية اختيار العينة يضمن عدم الانحياز لفئة دون أخرى أو نوع دون آخر. العشوائية هنا تعني استخدام الأسلوب العلمي وليس بالمفهوم المتعارف عليه في اتخاذ القرارات بدون دراسة مسبقة لها. إن العشوائية في الاختيار تتيح فرصا باحتمالات متساوية للجميع ولا يتم الانحياز لجانب دون آخر. ففي مسألة الإسكان في مدينة (س) من المدن لا ينحاز الباحث أثناء جمعه لبيانات للمساكن الفارهة ويهمل ألمساكن الفقيرة بل يجب أن تعطى كل المساكن فرص متساوية حتى تكون نتائج الدراسة ممثلة الواقع تمثيلا صحيحا لأبعد الحدود.

والأسلوب العلمي يبدأ بالملاحظة التي تثير في النفس الإحساس بالمشكلة ويترتب على هذا الإحساس وضع الفرض العلمي المبدئي لتفسير هذه المشكلة ثم السعي للتحقق من صحة هذا الفرض أو عدم صحته وذلك بإجراء الدراسة الميدانية أو المعملية لجمع البيانات وتحليلها للوصول إلى النتائج المفسرة. إن الأسلوب العلمي يتطلب دقة في اختيار المشكلات، ودقة في جمع البيانات ورصد النتائج وهذا لا يتأتى بطبيعة الحال إلا إذا كانت الدراسة على نطاق محدود تشمل منطقة صغيرة أو عددا محدودا من الناس (أسلوب العينة) حيت يكتفي الباحث بدراسة مجموعة ممثلة بدلا من دراسة جميع مناطق أو أفراد الظاهرة المطلوب دراستها لأن في هذا الأسلوب توفير للوقت والجهد والتكلفة كما سبقت الإشارة إلى ذلك. وإذا أيدت البيانات المجمعة وتحليلاتها المستفيضة الفرضية العلمية، فهذه الفرضية تعتبر تفسيرا للمشكلة أو حكما عليها. وإذا تكرر هذا التفسير أو الحكم مرارا وتكرارا فقد يسمى تعميما (generalization) أو قاعدة (low) أو نظرية (theory) وفى هذه الحالة يطبق على الظواهر والحالات المماثلة الماضية والحاضرة والمستقبلة. إن إيجاد النظريات أو التعميمات يمكن أن تستخدم في التنبؤ لحلول مشاكل جديدة على نسق

المشاكل السابقة، لذلك فان موضوع التنبؤ وتقديم الحلول أصبح من أهداف الجغرافيا الحديثة شأنها في ذلك شأن العلوم الأخرى.

ماذا يحصل عند مخالفة شَرطي كبر حجم العينة والعشوائية في اختيار مفرداتها؟

الهدف من الدراسات العلمية هو إصدار أحكام سليمة تخص مجتمع الظاهرة المدروسة وهذا لا يحصل إلا باستيفاء شرطين هما: (كبر حجم العينة وعشوائية اختيار مفرداتها) أما إذا كانت العينة صغيرة أو منحازة أو كليهما فأمامنا ثلاثة احتمالات هي:

1- أحكام مبالغ فيها مغالطة للواقع بنسبة 33% من المناسبات.

2- أحكام مبخسة للواقع بنسبة 33% . وإذا جمعنا التنبؤات المبالغة والمبخسة معا فنحن أمام 66% من الأحكام الخاطئة مما يعقد الأمور ولا يحل المشكلة.

3- أحكام جيدة مطابقة للواقع باحتمال 33% من المناسبات وهذا ناتجا عن الصدفة وليس لدقة التحليل.

المتغير (variable):

وهو يعني الظاهرة المدروسة كالذبذبات في المطر السنوي، والحرارة، ودرجات الطلبة في الامتحان ومدة الرحلة من البيت إلى مكان العمل ومبيعات ليبيا من النفط، وكميات الإنتاج الزراعي، وحوادث المرور والنمو السكاني وتنامي ظاهرة الهجرة غير الشرعية والاتجار في الممنوعات، إلى غير ذلك. وفي كثير من الأحيان نهتم بدراسة العلاقة بين متغيرين، كالعلاقة بين إنتاج القمح وكميات الأمطار، أو بين مرض السكر والوزن، وبين التفوق العلمي ومستوى الدخل، وحوادث المرور والسرعة، الخ. وفي هذه الأحوال وغيرها، يوجد لدينا متغيرا حرا يأخذ الرمز (x) وآخر معتمد عليه يأخذ الرمز (y). فالقمح معتمدا على المطر لذلك نرمز للقمح y وللمطر x. والحوادث معتمدة على السرعة لذلك نرمز للحوادث y وللسرعة x.

وعند توقيع البيانات على ورقة الرسم دائما ننصح بتوزيع قيم المتغير الحر x على المحور الأفقي وقيم المتغير المعتمد y على المحور الرأسي كما هو موضح في الأمثلة التالية:

مثال (1):-

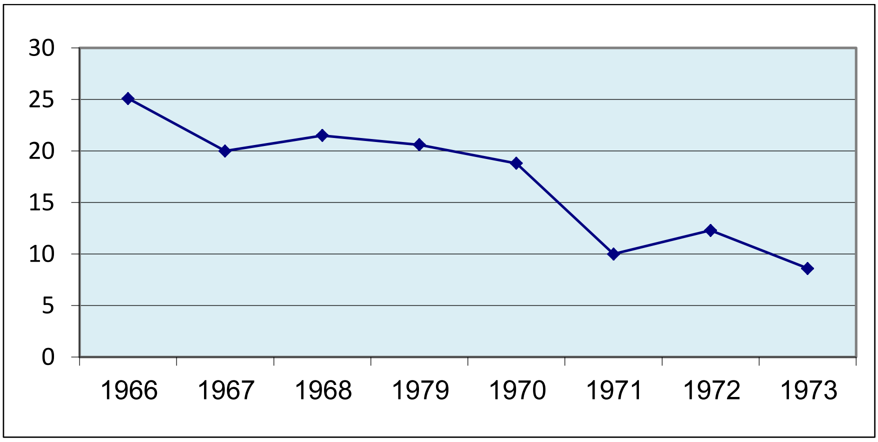
الجدول التالي يبين الصادرات من البضائع العامة من ميناء طرابلس خلال الفترة 1966-1973، والمطلوب رسم تخطيط الانتشار للسنوات على المحور الأفقي وكميات الصادرات على المحور الرأسي لمعرفة نوع العلاقة موجبة أم سالبة.

جدول (1) صادرات ميناء طرابلس للفترة 1966–1973 بآلاف الأطنان.

السنة	66	67	68	69	70	71	72	73
الكمية	25.1	20.0	21.5	20.6	18.9	10.0	12.3	8.6

المصدر: أبوالقاسم العزابي وفوزي الأسدي أبو صفحة وبشير أبوقيلة: دليل الباحث، ط/2، المنشأة العامة للنشر والتوزيع والإعلان، طرابلس، 1982.

شكل(1) صادرات ميناء طرابلس للفترة 1966 –1973 (ألف طن).



من الشكل يتبين إن اتجاه الصادرات من ميناء طرابلس سالب فبينما كان عام 1966 أكبر من 25 ألف تناقص إلى أقل من تسعة آلاف عام 1973. وعلى الباحث إجراء دراسة عن الأسباب التي أدت إلى هذا الوضع غير السليم في الاقتصاد الليبي.

مثال (2):-

الجدول التالي يبين واردات النفط عبر ميناء طرابلس خلال الفترة 1966 –1973، الكمية بالأطنان، والمطلوب رسم تخطيط الانتشار لمعرفة نوع الارتباط .

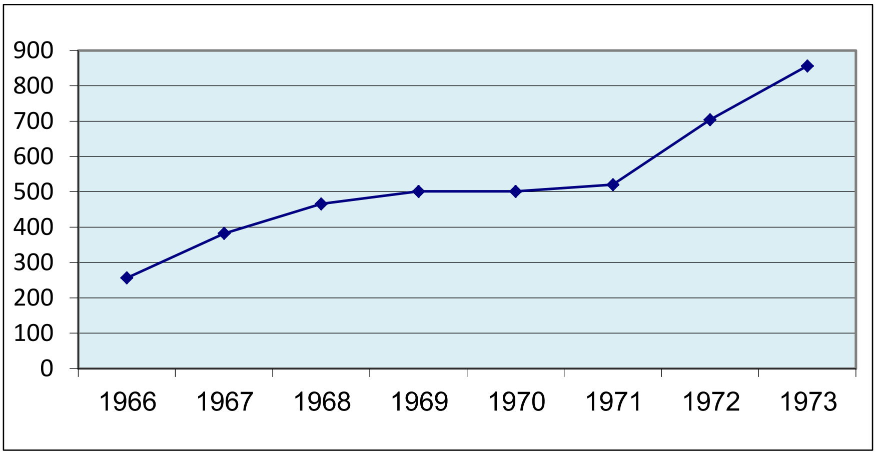
جدول (2) واردات النفط عبر ميناء طرابلس للفترة 1966 –1973(ألف طن).

السنة	66	67	68	69	70	71	72	73
الكمية	256	382	466	501	501	520	704	856

المصدر: أبوالقاسم العزابي وفوزي الأسدي أبو صفحة وبشير أبوقيلة: دليل الباحث، ط/2، المنشأة العامة للنشر والتوزيع والإعلان، طرابلس، 1982.

من الجدول رقم (2) والشكل (2) يتبين أن الوضع الخاص باستيراد النفط يأخذ شكلا مغايرا لتصدير المنتجات الليبية إلى الخارج هنا اتجاه استيراد النفط في زيادة سريعا سنة وراء سنة والأسباب معروفة لا تخفى على أحد.

شكل(2) واردات النفط عبر ميناء طرابلس للفترة 1966– 1973(ألف طن).



من شروط العينة الجيدة التجانس:

تهدف الدراسات العلمية بمختلف تخصصاتها إلى التحقق من خصائص الظواهر ونظم تغيرها في الزمان والمكان من أجل الوصول إلى تنبؤات دقيقة تفيد في التخطيط السليم الذي يضمن اتخاذ القرارات السليمة التي تقلل من الجهد والوقت المبذولان في سبيل إنجاز المشاريع الإنمائية كما تقلل أيضا من تكلفتها الإنشائية. تتعدد المشاريع التنموية فمنها المتعلق بالطرق والكباري ومنها خاص بالتعليم (كالمدارس والمعاهد والجامعات) ومنها ما يتعلق بالتوسع الزراعي والرعوي أو بالصحة(كالمصحات والعيادات الخاصة والمستشفيات والصيدليات). كل هذه المشاريع الحيوية تحتاج قبل البدء في إنجازها إلى إجراء دراسات متعلقة بأعداد السكان وتوزيعات الكثافات السكانية واتجاهات النمو السكانية والكثافات المرورية والأمراض والتغيرات الموسمية في المناخ إلى غير ذلك.

وأفضل وسيلة سريعة مستخدمة للتنبؤ المستقبلي بخصائص معا لم العينة هو المتوسط الحسابي المحسوب من قسمة مجموع قيم المشاهدات على عددها :

$$\overline{X} = \frac{\sum X}{n}$$

حيث أن : $\sum X$ مجموع قيم المشاهدات في العينة.

N = عدد مشاهدات العينة.

$\overline{X}$ = المتوسط الحسابي للعينة.

مثال: البيانات التالية هي للمتوسطات الشهرية للحرارة(مئوية) المسجلة بمدينة مسلاته خلال عام 1993 والمطلوب حساب المتوسط الحسابي لمجمل السنة.

جدول(3) المتوسطات الشهرية للحرارة بمدينة مسلاته خلال عام 1993.

يناير	فبراير	مارس	ابريل	مايو	يونيو	يوليو	أغسطس	سبتمبر	أكتوبر	نوفمبر	ديسمبر
12.9	11.9	14.9	17.5	20.9	24.7	26.9	26.6	26.1	24.5	20.0	15.3

$$\sum X = 242.2$$
$$\overline{X} = 242.2 \div 12 = 20.18\ C$$

لاحظ هنا أن المتوسط الحسابي للحرارة الشهرية في مسلاته(20.18) درجة مئوية غير جيد في التنبؤ لكل الشهور لذلك نقول بأن العينة غير متجانسة والدليل على ذلك أن أشهر ديسمبر ويناير وفبراير ومارس أقل منه بكثير وأن أشهر يوليو وأغسطس أكبر منه بكثير ولا يقترب منه إلا أشهر نوفمبر وأبريل ومايو ويونيو وأكتوبر. وسوف نعود للتفسير أكثر بعد قليل.

القدرة الشرائية للسكان تختلف باختلاف معدلات الدخل بين الأحياء الغنية والمتوسطة والفقيرة كما تختلف أذواق الزبائن لأنواع السلع حسب دخولهم لذلك فمن العبث إقامة معرض للسيارات الرياضية والفارهة بالأحياء الفقيرة أو فتح معارض للسيارات الشعبية الرخيصة بالأحياء الغنية. ونفس الشيء يمكن أن يقال عن محلات بيع الملابس، فالمستثمر لا بد أن يدرس معدلات الدخل بين الأحياء لتحديد نوع السلعة الممكن تصريفها وتحقق له الأرباح. ومن العبث عرض الملابس الخاصة بالرياضة البحرية في مدن صحراوية بعيدة عن البحر فلن يشتريها أحد ولا عرض ساعات ذهبية بآلاف الدولارات في أسواق الأحياء الفقيرة. لذلك فالتاجر الناصح مضطر لجمع بيانات عن الدخل السنوي للأسر المنتشرة بين الأحياء السكنية ومعرفة متوسطات الدخل من أجل اختيار المواقع المناسبة لاستثماراته.

والاستخدام الجيد للمتوسط الحسابي في التنبؤ يشترط تجانس مفردات العينة المدروسة (جميع المفردات من نفس المجتمع لذلك يجب أن تكون متقاربة في الحجم ولا يوجد تفاوت كبير بينها)، أما إذا كان هناك تباين كبير بين المفردات (لا يتحقق شرط التجانس) فلا يصلح عندئذ المتوسط كوسيلة للتنبؤ. تأمل المثال التالي لعينتين من المطر السنوي :

العينة الأولى متجانسة أو قريبة من التجانس:

جدول (4) بيانات المطر بالمحطة الأولى(ملم).

7	6	5	4	3	2	1
145	155	130	170	140	160	150

العينة الثانية **غير المتجانسة** فيها المفردة الرابعة التي تحتها خط خارجة عن النسق:

جدول (5) بيانات المطر بالمحطة الثانية(ملم).

7	6	5	4	3	2	1
145	155	130	**300**	140	160	150

ما معنى خارجة عن النسق؟: المفردة رقم 4 في العينة الثانية، جدول (5) كبيرة جدا ولا يمكن اعتبارها من نفس المجتمع وإذا بقت في المجموعة فإنها تؤدي إلى تضخم المتوسط الحسابي ويصبح غير ممثل تمثيلا جيدا للمجتمع المأخوذ منه ويقود استعماله في التنبؤ إلى نتائج خاطئة.

- كيف نعرف أن العينة **متجانسة أو غير متجانسة؟**

من خلال توزيع المفردات حول المتوسط الحسابي للعينة نستطيع الحكم على أن العينة متجانسة أو غير متجانسة.

الخطوات:

1- نحسب المتوسط الحسابي للعينة الأولي والعينة الثانية.

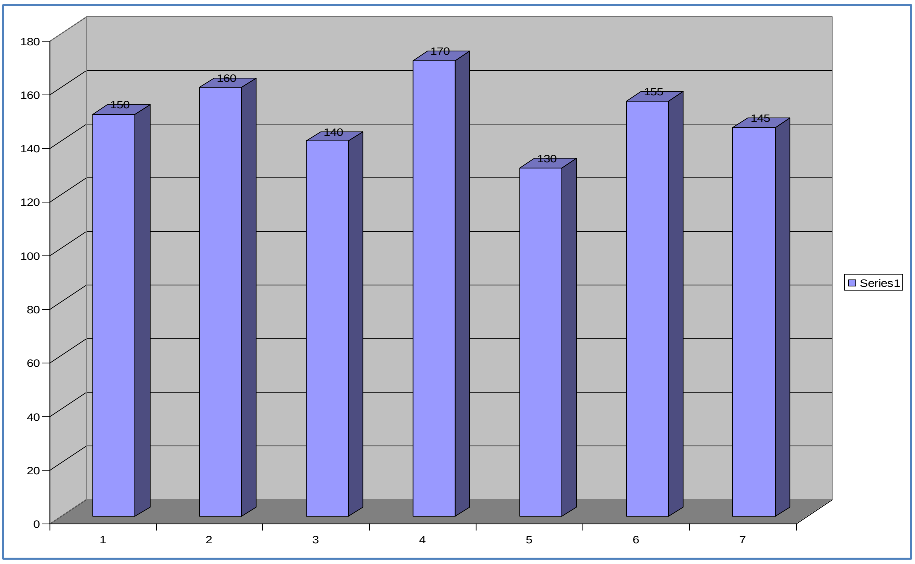
$$\overline{X}_1 = \frac{1050}{7} = 150$$

$$\overline{X}_2 = \frac{1080}{7} = 168.5$$

2- نقارن بين مفردات العينة الأولى وبين متوسطها الحسابي (150) فنجد أن القيم متقاربة مع المتوسط فهي إما مساوية للمتوسط كما في المفردة الأولى أو أكبر منه قليلا كما في المفردة الثانية والرابعة والسادسة أو أقل منه قليلا كما في المفردة الثالثة والخامسة والسابعة. لذلك نحكم بأن العينة متجانسة ويمكن استخدام الرقم 150 كوسيلة للتنبؤ للسنة اللاحقة فهي إما أن تساوي 150 أو أزيد منه أو أقل قليلا. أنظر الشكل رقم (3).

3- من مقارنة مفردات العينة الثانية بمتوسطها الحسابي (168.5) يتبين أن كل القيم ماعدا الرابعة أقل من المتوسط الحسابي فما هو السبب؟. في حالة وجود مفردة متطرفة الحجم (أنظر الشكل التالي) فإنها تؤدي إلى تضخم المتوسط الحسابي وتجعله ينحاز نحوها فيصبح متوسط مخادع لا يعبر عن حقيقة التوزيع ولا يمكن استخدامه كوسيلة للتنبؤ المستقبلي.

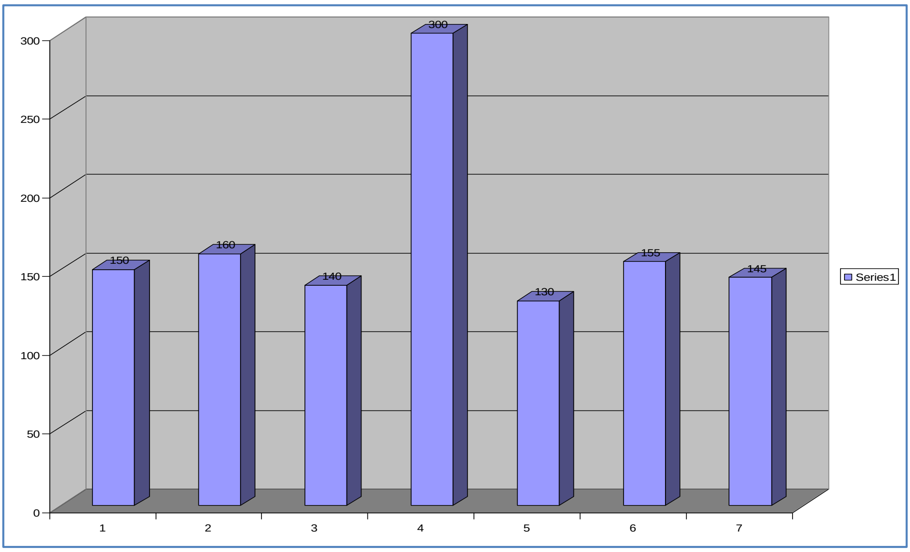
شكل (3) العينة متجانسة ويمكن استخدام المتوسط الحسابي في التنبؤ.



تعتمد مبيعات السلع الغذائية من لحوم وأسماك وألبان وخضروات على العلاقات التي يبنيها التاجر مع زبائنه، فإذا كان التاجر أمينا لا يعرض إلا السلع الطازجة وأسعاره معقولة تنبني علاقة ثقة بينه وبين زبائنه فلا يفضلونه على غيره فتبقى مبيعاته منتظمة ويستطيع أن يخطط لمشترياته لتي سيبيعها في الأسابيع القادمة. أما إذا كانت مبيعاته غير منتظمة (يوما يبيع بعشرة ويوما بألف) نتيجة غياب الثقة بينه وبين الزبائن فتجد أسعاره مبالغ فيها ويعرض سلعه للتلف نتيجة إطالة مدة العرض والتخزين مما يضطره إلى التخلص منها كقمامة. فهذا التاجر لا يستطيع أن يخطط لمشترياته

التي سيبيعها في الأسابيع القادمة لأن سجل مبيعاته غير منتظم، مما يعرضه إلى تكبد الخسائر نتيجة تلف السلع وتحول الزبائن إلى تجار آخرين. وهكذا.

شكل(4) العينة الثانية تحتوي على المفردة الرابعة خارجة عن النسق، مما ضخم المتوسط الحسابي وجعله غير صالح للتنبؤ(فالعينة غير متجانسة).



في جميع الأحوال يتبين أن مسألة التجانس في مفردات العينة المدروسة مهمة جدا. فالتجانس يعني أن تشتت المفردات حول المتوسط الحسابي قليل ويعني أن كل مفردات العينة تنتمي إلى نفس المجتمع قيد الدراسة ومن ثم فإن النتائج المتوصل إليها في الدراسة تكون حقيقية ذات مصداقية عالية ولا تشوبها شائبة.

وتعتمد أرباح خطوط النقل الجوي والبحري والسكك الحديدية وكذلك البرية على درجة انتظام رحلاتها، فالشركات المشهورة بدقة انتظام مواعيدها في نقل الركاب والبضائع تنال ثقة الزبائن وتنال سمعة جيدة مما يؤدي إلى تفضيلها على غيرها من شركات النقل وهذا يؤدي بالطبع إلى تنامي أرباحها. أما الشركات التي تتصف بتذبذب مواعيد رحلاتها مرات بالتأخير وأخرى بالتبكير مما يسبب مشاكل للمسافرين الذين يتأخرون عن أعمالهم الرسمية ويعرضهم إلى عقوبات نتيجة ذلك التأخير فإنهم يلجئون إلى شركات أخرى أكثر مصداقية في مواعيدها، مما يضطر الشركات ذات المصداقية القليلة إلى تسيير رحلاتها بنسبة أقل من الحمولة الاقتصادية المربحة ويعرضها إلى الخسائر وربما الإفلاس.

المدى Range:

يحسب المدى R بطرح أصغر القيم من أكبر القيم في العينة. ويستخدم عند المقارنة بين توزيعين لمعرفة أيهما أكثر تجانسا وأيهما أكثر تشتتا في مفرداته. فالتوزيع المتجانس يحتوي على مفردات متقاربة في الحجم لذلك يصغر مداه ، أما التوزيع غير المتجانس فيظهر تباعدا كبيرا بين المفردات ويظهر ذلك في كبر المدى.

استخدم بيانات المثال السابق في حساب المدى لكل عينة ثم حدد أيهما أكثر تجانسا وأيهما أقل تجانسا في مفرداته واشرح كيف توصلت إلى تلك النتيجة.

1- حساب المدى للعينة الأولى:

لحساب المدى نوجد أكبر القيم وأصغر القيم في العينة.

أكبر القيم=170

أصغر القيم=130

المدىR= أكبر القيم– أصغر القيم

= 170 - 130=40

2- حساب المدى للعينة الثانية :

أكبر القيم=300

أصغر القيم=130

R=300-130=170

المقارنة والاستنتاج:

من مقارنة قيمة المدى للعينة الأولى (40) والمدى للعينة الثانية (170) يتبين أن العينة الأولى أكثر تجانسا لصغر مداها، وأن العينة الثانية أقل تجانسا لكبر مداها. فالعلاقة بين المدى والتجانس علاقة عكسية فكلما كبر المدى قل التجانس وكلما صغر المدى زاد التجانس. وبتعبير آخر العينة صغيرة المدى مفرداتها قريبة من متوسطها الحسابي بينما العينة كبيرة المدى مفرداتها متباعدة وكبيرة التشتت حول متوسطها الحسابي.

الانحراف المعياري:

المدى معامل يسهل حسابه ويعطي فكرة سريعة عن التشتت في قيم العينة إلا أن مشكلته أنه يحسب من مفردين فقط من مفردات العينة ويتجاهل الباقي لذلك لا يستعمل في إجراء اختبارات الدلالة الإحصائية ويستخدم بدلا منه معامل الانحراف المعياري الذي يعتمد في حسابه على انحرافات جميع القيم عن متوسطها الحسابي كما يلي:

$$S = \sqrt{\frac{\sum (X - \overline{X})^2}{n-1}}$$

حيث أن : S= الانحراف المعياري للعينة.

$\sum (X - \overline{X})^2$ = مجموع تربيع انحرافات القيم عن المتوسط الحسابي.

N= حجم العينة(عدد المفردات أو القيم بالعينة).

مثال: استخدم بيانات المطر من المثال السابق لحساب الانحراف المعياري للعينتين وحدد أيهما أكثر تجانسا في مفرداته وبين السبب في وصولك لتلك النتيجة؟

جدول (4) بيانات المطر بالمحطة الأولى(ملم).

7	6	5	4	3	2	1
145	155	130	170	140	160	150

جدول (5) بيانات المطر بالمحطة الثانية(ملم).

7	6	5	4	3	2	1
145	155	130	300	140	160	150

الحل:

من التمرين السابق عرفنا أن المتوسط الحسابي للعينة الأولى هو 150 ملم والمتوسط الحسابي للعينة الثانية هو 168.8 ولحساب الانحراف المعياري لكل عينة نكون جدول نحسب فيه الانحراف التربيعي للقيم عن المتوسط ثم نحسب مجموع تلك الانحرافات كما يلي:

جدول(6) حساب معطيات الانحراف المعياري

للعينتين الأولى والثانية

الانحراف المعياري للعينة الأولى		الانحراف المعياري للعينة الثانية	
X_1	$(X-\overline{X})^2$	X_2	$(X-\overline{X})^2$
150	0	150	342.25
160	100	160	72.25
140	100	140	812.25
170	400	300	17292.25
130	400	130	1482.25
155	25	155	182.25
145	25	145	552.25
$\sum=1050$	$\sum=1050$	$\sum=1180$	$\sum=20735.75$

أولاً: معطيات العينة الأولى:

$$\sum(X-\overline{X})^2=1050$$

N=7

$$\sum X1=1050$$

$$\overline{X}=150$$

نعوض في المعادلة ونحسب قيمة الانحراف المعياري الأول كما يلي:

$$S1=\sqrt{\frac{1050}{7-1}}=13.22mm$$

إذا قيمة الانحراف المعياري= 13.22 مليمتر.

ثانياً: معطيات العينة الثانية:

$$\sum(X-\overline{X})^2=20735.75$$

N=7

$$\sum X2=1180$$

$$\overline{X}=168.57$$

نعوض في المعادلة ونحسب قيمة الانحراف المعياري الثاني كما يلي:

$$S2 = \sqrt{\frac{20735.75}{7-1}} = 58.78mm$$

أذا قيمة الانحراف المعياري الثاني= 58.78 مليمتر.

المقارنة والاستنتاج:

من مقارنة قيمة الانحراف المعياري للعينتين يتبين أن قيمة الانحراف المعياري للعينة الأولى أصغر مما يعني أنها أكثر تجانسا وانتظاما في مطرها. فالعلاقة دائما عكسية بين حجم الانحراف المعياري والتجانس، فكلما قلت قيمة الانحراف المعياري زاد التجانس في العينة، وبالعكس يقل التجانس مع زيادة حجم الانحراف المعياري.

مما تقدم يتبين أن استخدام المتوسط الحسابي للعينة الأولى في عمل تنبؤات مستقبلية بخصوص المطر سيوصلنا إلى نتائج أفضل وأقرب للواقع وهامش الخطأ بها صغيرا مقارنة بالنتائج الممكن الحصول عليها باستخدام متوسط العينة الثانية لعدة أسباب وهي:

1- مفردات العينة الأولى قريبة من المتوسط ومنتشرة حوله بانتظام تقريبا فهناك ثلاثة مفردات فوق المتوسط وثلاثة تحته، أما مفردات العينة الثانية فكلها ما عدا واحدة أصغر من المتوسط الحسابي للعينة مما يدل على عدم التجانس في توزيع مفردات العينة.

2- المدى في العينة الأولى أصغر وكما هو معروف أن العلاقة عكسية بين المدى والتجانس.

3- الانحراف المعياري للعينة الأولى أصغر لذلك فالعينة الأولى أكثر تجانسا، والتجانس شرط ضروري من شروط عمل تنبؤات مستقبلية سليمة.

معادلة الانحراف المعياري السابقة تستخدم في أغلب الأحوال مع المفردات الكبيرة (مئات وآلاف) ، أما في حالة المفردات الصغيرة (عشرات وكسور) يفضل استخدام المعادلة التالية، علما بأن استخدام أيا منها سيعطى نفس النتائج و لك الحرية في اختيار التي تراها مناسبة.

$$S = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N-1}}$$

حيث أن: $\sum X^{2}$ = مجموع مربع القيم.

$(\sum X)^{2}$ = مجموع القيم الكل تربيع.

مثال:

الجدول التالي يبين أعداد الأحذية المباعة في أحد المتاجر خلال عشرة أيام والمطلوب استخدام معادلة الأرقام الصغيرة لحساب الانحراف المعياري لمبيعات الأحذية.

جدول (7) مبيعات أحد المتاجر من الأحذية خلال 10 أيام.

اليوم	عدد الأحذية	اليوم	عدد الأحذية
1	7	6	13
2	10	7	7
3	6	8	11
4	12	9	5
5	9	10	8

الحل: نكون جدول نحسب فيه عناصر معادلة الانحراف المعياري كما يلي:

X^2	X
49	7
100	10
36	6
144	12
81	9
169	13
49	7
121	11
25	5
64	8
المجموع =838	**المجموع =88**

المعطيات:

$N=10$

$\sum X^2 = 838$

$\sum X=88$

نعوض في المعادلة ونحسب قيمة الانحراف المعياري كما يلي:

$$S = \sqrt{\frac{838 - \frac{(88)^2}{10}}{10-1}} = 2.65$$

الانحراف المعياري لمبيعات الأحذية = 2.65 حذاء في اليوم.

ماذا يعني الانحراف المعياري؟

في دراسات العينة نفترض أن التوزيع التكراري لمفردات العينة يأخذ التوزيع الاعتدالي (الجرسي) ذو الخصائص التالية:

1- المنحنى اعتدالي التوزيع يعنى أن المتوسط يقسم العينة المرتبة ترتيبا تصاعديا إلى 50% من المفردات أصغر منه و 50% من المفردات أكبر منه. وفي مثال الأحذية 50% من الأيام مبيعاتها ستكون أكبر 8.8 حذاء في اليوم و50% ستكون مبيعاتها أصغر من 8.8 حذاء في اليوم.

2- 68% من مبيعات التاجر ستكون في حدود المتوسط + أو – واحد انحراف معياري وبما أن المتوسط= 88÷10=8.8 حذاء فمبيعاته ستكون بين 8.8+ 2.65 =11.46 و 8.8-2.65= 6.15 حذاء في اليوم، أي بين 11.46 و 6.15 حذاء في اليوم.

3- 95% من مفردات العينة تتوزع حول المتوسط الحسابي مضافا إليها إنحرافين معياريين على اليمين و المتوسط – إنحرافين معياريين من على اليسار أي: 8.8+2(2.65)=14.04 حذاء و 3.5 حذاء في اليوم.

معامل التباين النسبي : (CV%) Coefficient of Variation

إن استخدام الانحراف المعياري في المقارنة بين عينتين يشترط أن تكون **العينتان من نفس النوع وبنفس الوحدات القياسية وبنفس الإقليم الجغرافي**، لذلك فاستعماله يكون غير مجدي في الحالات التالية:

1- عند مقارنة عينتين من نفس النوع لكن في إقليمين متباعدين جغرافيا، كالمقارنة بين المطر في طرابلس والمطر في هون (هنا الموقعان مختلفان مناخيا الأول تابع للمناخ شبه البحر المتوسطي والثاني تابع للمناخ الصحراوي)، أو عند المقارنة بين معدلات إنتاج القمح بين ليبيا وتركيا أو بين معدلات إنتاج الأرز بين الصين والولايات المتحدة. فبالرغم من أن السلعة واحدة إلا أن الظروف المتحكمة في الإنتاج مختلفة بين الإقليمين، واستخدامنا للانحراف المعياري كأساس للمقارنة يقودنا إلى نتائج مغالطة.

2- عند مقارنة عينتين لنفس النوع في نفس الإقليم لكن بمقياس مختلف كمقارنة عينتين من الأرز الأولى بالقناطير والثانية بالطن المتري، أو الأولى بالأرطال والثانية بالكيلوات.

3- عند مقارنة نوعين مختلفين كالمقارنة بين إنتاج النفط والفحم، أو بين صادرات الأقمشة وبين صادرات الأحذية.

والسؤال المطروح ما هو البديل الممكن استخدامه في تلك الحالات التي لا ينفع معها استخدام الانحراف المعياري كأساس للمقارنة؟ البديل الفعال هو معامل الاختلاف النسبي (CV% ، Coefficient of Variation الذي يحسب بقسمة الانحراف المعياري على المتوسط الحسابي ويضرب الناتج في 100.

وفي الوقت الذي يحمل فيه الانحراف المعياري والمتوسط الحسابي والمدى نفس الوحدات القياسية لمفردات العينة فإذا كانت القيم قناطير كان المتوسط والانحراف المعياري والمدى قناطير كذلك وإذا كانت القيم ملليمترات كان المتوسط والانحراف المعياري والمدى ملليمترات كذلك إلا أن معامل التباين النسبي CV عبارة عن نسبة لا وحدات لها، لذلك يمكن استعماله كأساس للمقارنة في كل الأحوال. ويمكن حساب CV بالمعادلة التالية:

$$CV = \frac{S}{\overline{\overline{X}}}100$$

مثال1- الجدول التالي يبين إنتاج القمح في إيران وبريطانيا خلال ثمانية سنوات والمطلوب إيجاد أي الدولتين أكثر انتظاما في إنتاجها:

جدول (8) إنتاج القمح في إيران وبريطانيا
خلال ثمان سنوات.

السنة	1	2	3	4	5	6	7	8
إيران	199	200	180	300	290	177	350	420
بريطانيا	109	120	130	140	137	125	133	118

نتائج التحليل:

الدولة	أكبر القيم	أصغر القيم	المدى R	$\overline{X}$	S	(CV%)
إيران	420	177	234	278.22	93.67	33.66
بريطانيا	172	109	63	130.33	17.79	13.64

نظرا لأن بريطانيا وإيران من إقليمين جغرافيين مختلفين لا نستطيع استخدام الانحراف المعياري للحكم على أيهما أكثر تجانسا في الإنتاج والأفضل استخدام معامل التباين النسبي، ومن مقارنة هذا المعامل للبلدين يتبين أن بريطانيا أكثر تجانسا في إنتاج القمح. لماذا؟.

مثال (2):- الجدول التالي يبين صادرات نيجيريا من الحديد والغاز الطبيعي والمطلوب إيجاد أي السلعتين أكثر انتظاما في التصدير.

جدول (9) صادرات نيجيريا من الحديد والغاز الطبيعي خلال 10 سنوات.

السنة	1	2	3	4	5	6	7	8	9	10
الحديد مليون (طن)	140	133	210	120	136	118	115	99	88	95
الغاز مليون (م^3)	316	218	150	100	217	360	260	144	117	220

نتائج التحليل:

السلعة	أكبر القيم	أصغر القيم	المدى R	$\overline{X}$	S	(CV%)
الحديد مليون (طن)	360	100	260	210.21	85.08	40.47
الغاز مليون (م^3)	210	88	122	125.41	34.56	27.55

نظرا لأن الحديد والغاز سلعتين مختلفتين في النوع وفي وحدة القياس لا نستطيع استخدام الانحراف المعياري كأساس للمقارنة والأفضل منه هو معامل التباين النسبي، ومن مقارنة هذا المعامل للسلعتين يتبين أن الغاز أكثر انتظاما في التصدير. لماذا؟.

أسئلة المراجعة:

1- اشرح المقصود بالمشكلة البحثية.

2- ما المقصود بالفرضية العلمية؟ أعط أمثلة للمشاكل التالية:

أ- تأخر سن الزواج.

ب- الازدحام المروري على الطرقات الليبية.

ج- تدني مستوى التحصيل العلمي بالمدارس الليبية.

د- تكدس القمامة على نواصي الطرقات.

هـ- ارتفاع أسعار السلع في الأسواق.

و- تدني الأداء في خدمة المواطنين في الإدارات العامة مقارنة بحسن الأداء في الشركات الخاصة.

3- اشرح ماذا ينتج عن القرارات العشوائية من احتمالات.

4- اشرح مسار التفكير العلمي عند إجراء البحوث.

5- بين لماذا تجرى البحوث على العينة وليس على المجتمع المأخوذة منه؟.

6- بين ماذا يحصل من مشاكل عند مخالفة شرطي كبر حجم العينة وعدم توفر العشوائية في اختيار المفردات؟.

7- اشرح الأسلوب المتبع في اختبار تجانس العينات إذا كانت ليست من نفس النوع.

الفصل الثاني

التوزيع الاعتدالي واستعمالاته

المقصود بالتوزيع الاعتدالي وخواصه:

يشتق توزيع الاحتمالات لغالبية إحصاءات العينة من التوزيع الاعتدالي Normal Distribution الذي يعتبر أهم توزيع احتمالي في النظرية الإحصائية. فإذا كان التوزيع التكراري للمتغير x اعتدالي فالشكل العام له يكون كالجرس المقلوب، متماثل التوزيع حول المتوسط الحسابي، ويعتمد على معلمين هما المتوسط الحسابي $\overline{X}$ والانحراف المعياري S ويمكن تلخيص أهم خواصه (شكل5) في النقاط التالية:

1- قيم المتوسط الحسابي $\overline{X}$ والوسيط (md) والمنوال (mo) متساوية .

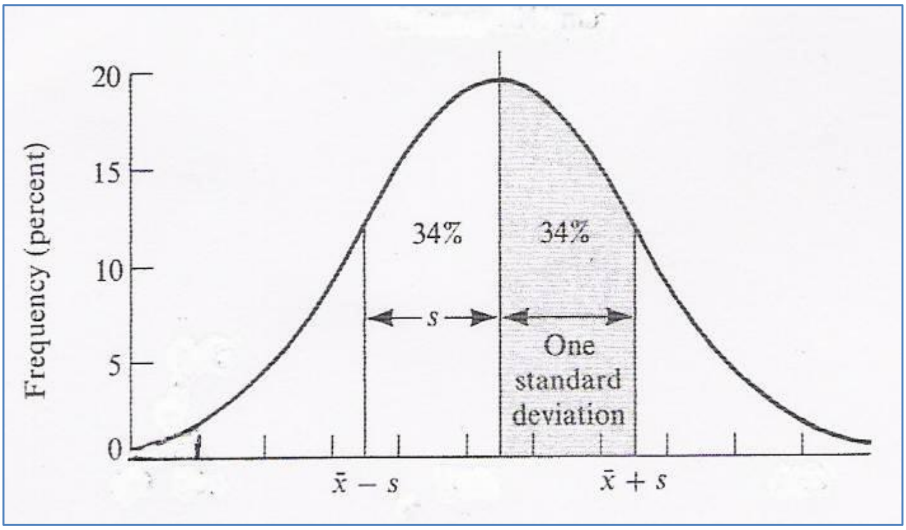
2- منحنى متماثل التوزيع يرتفع عند الوسط وينخفض تدريجيا نحو الطرفين ولا يلمس المحور السيني أبداً. والطرف الأيمن ينطبق تماما على الأيسر وكلاهما يشكل 50% من المساحة أو الاحتمالات تحت المنحنى .

3- المساحة الواقعة بين + S و – S تحتوي على 68.27 % من القيم والمساحة الواقعة بين +S2 و – S2 تحتوي على 95.45 % أما المساحة الواقعة بين +S3 و – S3 فهي تحتوي على 99.73% .

4- نقطتي تحول المنحنى، أي النقطتين اللتين يبدأ فيهما المنحنى يغير اتجاهه تقابلان +S و-S.

عموما إذا كان التوزيع قليل الالتواء أي قريب من التوزيع ألاعتدالي فانه يبقى محتفظا تقريبا على هذه الخواص الأربعة.

شكل(5) العلاقة بين التكرار ونسبة المساحة التي يمثلها تحت المنحنى الاعتدالي ما بين ± انحراف معياري من المتوسط يقع 68% من المشاهدات الممثلة بهذا التوزيع.



استعمالات المنحنى الاعتدالي:

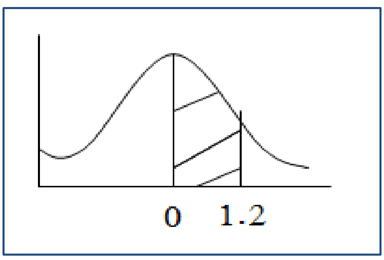
يستخدم المنحنى ألاعتدالي في إيجاد الاحتمالات في شكل نسب مئوية تحت المنحنى. لاحظ أن مجموع المساحة تحت المنحنى = 1 مقسمة إلى 0.50 أي 50% على الطرف الأيمن و0.50 أو 50% على الطرف الأيسر من المتوسط الحسابي الذي هو 0. وموقع كل مساحة تبينه قيمة Z والمتوسط الحسابي . فإذا كانت Z موجبة أبحث في الطرف الأيمن وإذا كانت سالبة فابحث في الطرف الأيسر.

أمثلة:

1- أوجد احتمال (النسبة المئوية تحت المنحنى) للحالات التالية:

أ - الدرجة المعيارية Z=1.2

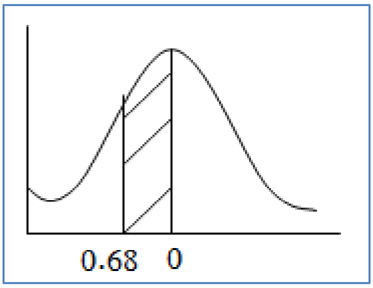
الحل: نرسم المساحة المطلوبة تحت المنحنى ونظللها كما يلي:



لاحظ أن المساحة المطلوبة واقعة بين المتوسط الحسابي أي Z=0 و Z =1.2 ، الآن نفتح جدول الاحتمالات ونحدد الرقم المقابل لقيمة 1.2 فنجده يساوي 0.3849 ويقرأ كما يلي : المساحة الواقعة بين المتوسط والدرجة المعيارية 1.2 تحتوي على 38.49 % من الحالات.

ب- المساحة الواقعة بين الدرجتين المعياريتين 0 و -0.68 .

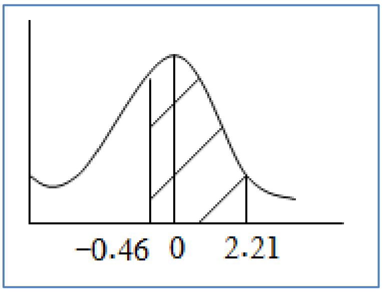
الحل: نحدد الموقع تحت المنحنى ثم نبحث في الجدول عن مساحته:



ومن الجدول المعياري يتبين أن هذه المساحة تشكل 0.2517 أي 25.17 % من الحالات.

ج- المساحة الواقعة بين الدرجتين - 0.46 و 2.21 .

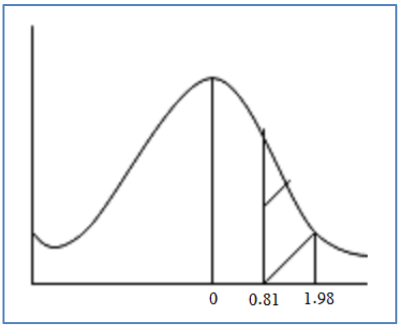
الحل: نحدد الموقع تحت المنحنى:



من الشكل يتبين أن هناك مساحتين يجب البحث عن قيمتيهما الأولى على يسار المتوسط وقيمتها من الجدول = 0.1772 والثانية على يمينه وتساوي 0.4864 ومجموع المساحتين = 0.6636 وتقرأ كما يلي : يوجد 66.36% من الحالات واقعة بين الدرجتين -0.46 و 2.21 .

د- المساحة الواقعة بين الدرجتين 0.81 و 1.98 .

الحل : نحدد الموقع تحت المنحنى :

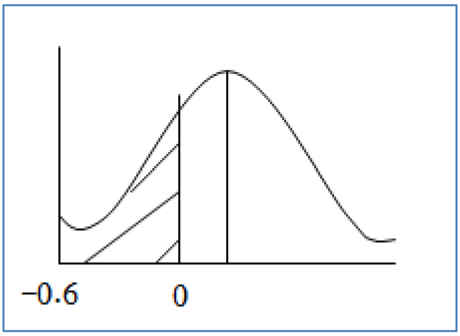


من الجدول يتبين أن المساحة بين 0 و 0.81 تساوي 0.2910 ، والمساحة بين 0 و1.98 تساوي 0.4761 ، لكن المطلوب حسابه هو المساحة بين 0.81 و 1.98 لذلك نطرح المساحة الصغرى من الكبرى كما يلي:

0.4761 -0.2951 = 0.1851 وهي تعادل 18.51% من الحالات.

هـ- المساحة الواقعة إلى يسار الدرجة 0.6 – .

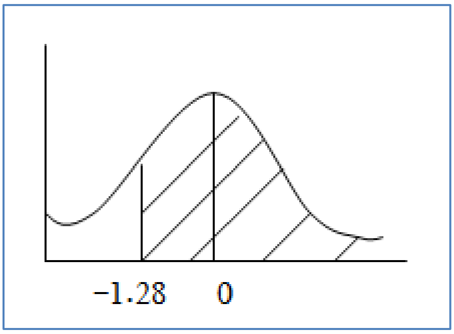
الحل: نحدد الموقع تحت المنحنى:



من الشكل يتبين أن المساحة المطلوب حساب احتمالها تقع إلى يسار الدرجة 0.6 – ومن الجدول يتبين أن هذه المساحة تشغل 27.4 %.

و – المساحة الواقعة إلى يمين 1.28 - .

الحل: نحدد الموقع تحت المنحنى:



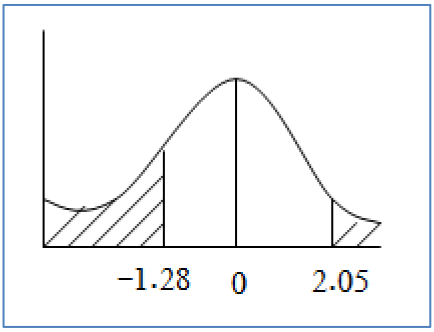
1- المساحة بين 1.28 و 0 تساوي 0.3997 والمساحة الواقعة إلى يمين 0 تساوي 0.50 .

2 – المطلوب كل المساحة الواقعة على يمين 1.28 – لذلك نجمع:

0.50+0.3997 = 0.8997 أي 89.98 % من الحالات .

ز- مجموع المساحتين الواقعتين إلى يمين الدرجة 2.05 والى يسار 1.44 - .

الحل: كما تعلمنا نحدد الموقعين تحت المنحنى ليسهل إدراك المطلوب في السؤال:



من الجدول المعياري نجد أن المساحة الواقعة على يسار 1.28 - =0.0749 أما المساحة الواقعة إلى يمين 2.05 فهي تساوي 0.0202 ومجموع المساحتين = 0.0951 أي 9.51% .

مثال 2- يبلغ متوسط أوزان 500 طالب 70 كيلوجرام وانحرافهم المعياري 7 كيلوجرامات، أفرض أن الأوزان موزعة اعتداليا، أوجد ما يلي:

أ- نسبة وعدد الطلبة الواقعة أوزانهم بين 50 و 75 كيلوجرام.

ب- نسبة وعدد الطلبة الأثقل من 77 كج.

ج- نسبة الطلبة الأقل من 60 كج.

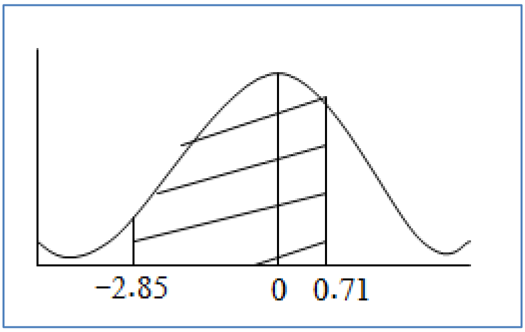
الحل: أ- نحول الأوزان إلى درجات معيارية بالمعادلة التالية:

$$Z = \frac{X - \overline{X}}{S}$$

$$Z_{50} = \frac{50-70}{7} = -2.85$$

$$Z_{75} = \frac{75-70}{7} = 0.71$$

نحدد موقع الدرجتين تحت المنحنى:



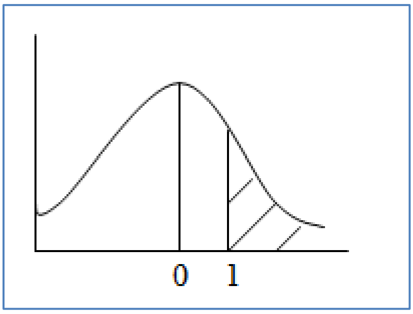
من الشكل يتبين أن هناك مساحتين يتوجب إيجاد احتمال كلا منهما من الجدول الأولى على يمين 0 والثانية على يساره. احتمال الأولى = 0.2611 أما احتمال الثانية =0.4978 ومجموع المساحتين =0.7589 أي 75,89% من الحالات . أما بخصوص عدد الطلبة الدين تشكلهم هذه النسبة فيمكن الحصول عليه بضرب النسبة في المجموع الكلي (النسبة × المجموع الكلي):

عدد الذين تقع أوزانهم بين 50 و 75 كيلوجرام =0.7589 ×500 =379 طالب .

ب- نسبة الطلبة الأثقل من 77 كج:

الحل: نحسب الدرجة المعيارية ثم نحدد الموقع تحت المنحنى ونقدر النسبة من الجدول ثم نحول النسبة إلى عدد:

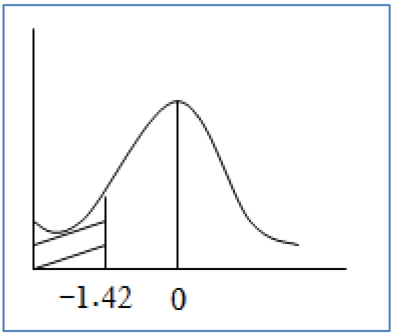
$$Z_{77} = \frac{77-70}{7} = 1$$



احتمال المنطقة الواقعة على يمين 1 =0.1587 أو 15.87% أما عدد الذين ينتمون لهذا الوزن = 0.1587 × 500 =79 طالب.

هـ- نسبة الطلبة الأقل وزنا من 60 كج:

$$Z_{60} = \frac{60-70}{7} = -1.42$$



أما احتمال هذه المجموعة فهو صغير إذ لا يشكل إلا 7.78 % فقط .

تمرينات:

1- أجري مسح لعدد 1000 منزل في المدينة وتبين أن متوسط أسعارها = 120 ألف دينار وانحرافها المعياري =15 ألف دينار، أوجد ما يلي:

أ- عدد المنازل التي تزيد أسعارها عن 140 ألف دينار.

ب- عدد المنازل التي تقل أسعارها عن 90 ألف دينار.

ج- عدد المنازل التي أسعارها أكبر من 100 وأقل من 130 ألف.

2- يمكن اعتبار أكياس الدقيق متغير عشوائي موزع اعتداليا بمتوسط 50 كج وانحراف معياري 1.5 كج . افرض انه تم أخذ أحد الأكياس عشوائيا، أوجد احتمالية احتواء هذا الكيس على:

أكثر من 52 كج.

أقل من 49 كج.

ج- بين 46 و 49 كج.

3- عدد الرسائل التي استلمها ساعي البريد يوميا ذات توزيع شبه اعتدالي بمتوسط يومي يساوي 500 رسالة وانحراف معياري 30 رسالة، أوجد احتمالية استلام:

أقل من 440 رسالة.

بين 530 و 575 رسالة.

ج- بين 440 و 470 رسالة.

4- إن كمية المطر السنوي في إقليم ما معروفة مسبقا بأنها متغير عشوائي يتبع التوزيع ألاعتدالي بمتوسط 125 سم وانحراف معياري 20 سم. وإذا تجاوزت كمية المطر في أي سنة من السنوات 135 سم فإنها تؤدي إلى حدوث فيضان مدمر، أوجد احتمالية حدوث هذا الفيضان لأي سنة من السنوات القادمة.

5- بلغ متوسط أسعار 760 سيارة تم بيعها في أحد السنوات 14800 دينار بانحراف معياري 2630 دينار . افرض أن الأسعار موزعة اعتداليا ، أوجد عدد السيارات التي تم بيعها بأسعار لم تتجاوز 15500 دينار .

6- افرض أن التوزيع التكراري للمطر السنوي بمدينة غريان متمشيا مع التوزيع ألاعتدالي بمتوسط 374.3 ملم وانحراف معياري 148 ملم ، أوجد ما يلي :

احتمالية سقوط أمطار أكبر من 500 ملم .

احتمالية سقوط أمطار أقل من 200 ملم .

التحليل التكراري بأسلوب المنحنى التراكمي:

يوجد أسلوب آخر للتحليل التكراري أسهل في التطبيق ويعطي نتائج مشابهة للتحليل التكراري السابق الذكر، فهو يستفيد من نظرية الاحتمالات لحساب المتوسط والانحراف المعياري ولتقدير النسب المئوية لتكرار أية قيمة يراد معرفتها. فالمجاميع السنوية للمطر في الأقاليم المطيرة تأخذ في الغالب التوزيع التكراري الاعتدالي (الشكل الجرسي) ذو الخصائص المعروفة.

والنموذج التالي يوفر الكثير من الجهد ويقدم الكثير من الخدمات و به يمكن تقدير قيم المتوسط والانحراف المعياري للعينة المدروسة واحتمالات أي قيمة يمكن أن تحدث في المستقبل. وفي هذا الأسلوب توقع القيم ضد (التكرار كنسبة مئوية) للحالات المساوية لهذه القيم أو أقل منها.

ونحصل على قيمة التكرار التراكمي بترتيب القيم تصاعديا بناء على أحجامها بحيث نعطي أصغر القيم الرتبة 1 بينما تتحصل أكبر القيم على الرتبة n.

وتقدر نسب التكرار لقيم المطر في هذه الحالة التي رتبتها m أو أقل منها بالمعادلة التالية:

$$F = \frac{m}{n+1} \times 100$$

وتفيد النسب التكرارية لتساقطات السجل المطري في التنبؤات المستقبلية فهي تستخدم كاحتمالات لتساقطات المستقبل أيضا. ويمكن رسم التوزيع التكراري التراكمي للمطر السنوي بتوقيع نسب التكرار F المحسوبة ضد قيم المطر المقابلة لها كما هو موضح بالشكل رقم (6) المحسوب لمدينة بالتيمور بولاية ماريلاند، حيث استخدمت ورقة رسم حسابية التوزيع على البعد الأفقي، واحتمالية التوزيع على البعد الرأسي. إن تقسيم المحور الرأسي بهذا الشكل يجعل المنحنى التكراري يأخذ شكل الخط المستقيم دلالة على أن القيم السنوية متمشية مع التوزيع الاعتدالي ذو الخصائص المتعارف عليها رياضيا.

ومن التوزيع التكراري التراكمي يمكن تقدير قيم المتوسط الحسابي والانحراف المعياري وقيمة أي احتمال آخر يرغب فيه كما يلي:

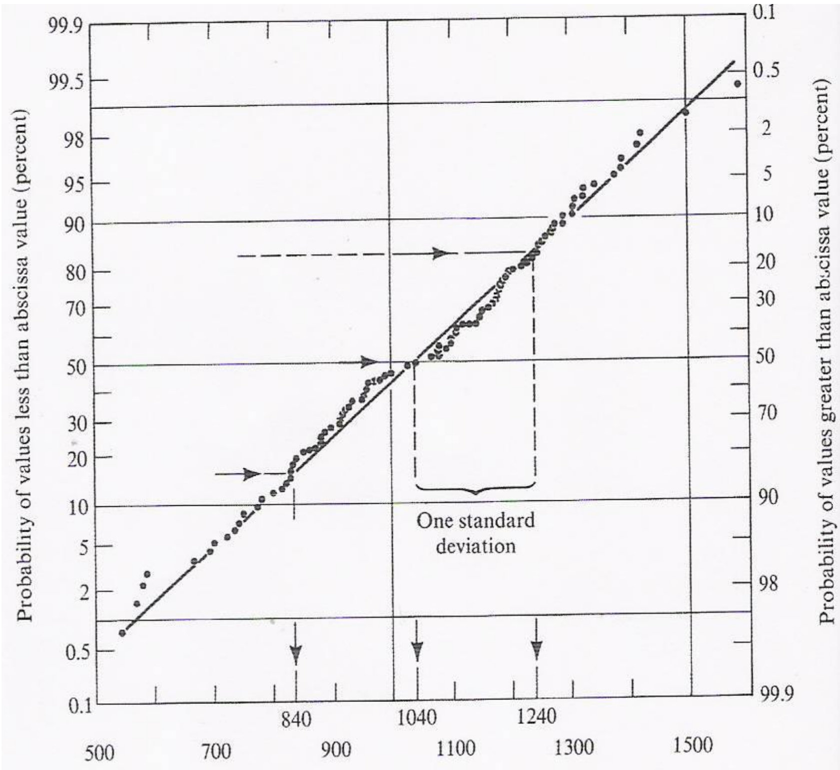
1- تقدير قيمة المتوسط الحسابي للتوزيع: لاحظ أن الاحتمالات موزعة على المحور الرأسي بينما قيم المطر (ملم) موزعة على المحور الأفقي، لذلك من احتمال 50 في المائة نتحرك في خط مستقيم نحو اليمين إلى أن نقطع خط الانحدار تم إلى أسفل فنجد أن قيمة المتوسط الحسابي تساوي 1040 ملم.

2- تقدير قيمة الانحراف المعياري: من المعروف أن 68% من كل المشاهدات تقع في المساحة الواقعة بين ± انحراف معياري من المتوسط الحسابي أي بين احتمالات 16 و 84 في المائة وبالتحرك من هاذين الاحتمالين نحو خط الانحدار ثم إلى أسفل نجد قيم المطر التي تتناسب مع انحراف معياري من كلا جانبي المتوسط كما يلي:

الأول بين 16 و 50 في المائة وتساوي 1040- 840=200 ملم.

والثاني بين 84 و 50 في المائة وتساوي 1240-1040=200 ملم.

شكل (6) التوزيع التكراري التراكمي للمطر السنوي بمدينة بالتيمور نيويورك.



المطـر (ملم)

ويمكن استخدام المنحنى التكراري التراكمي أيضا في تقدير احتمالات سقوط أمطار سنوية مساوية لأي قيمة سنوية أو أقل منها. فعلى سبيل المثال، هناك احتمال قدره 4.7 في المائة لأي سنة في مدينة بالتيمور لأن تكون أقل من 700 ملم، وهذا يعني أن أغلب الأمطار السنوية بواقع 95.3% هي أكبر من 700 ملم. وتفيد مثل هذه النتائج في التصميم لتمديدات المياه ومخططات التنمية الزراعية والعمرانية والصناعية. أما إذا كان المخطط مهتم بتصريف الأراضي أو غيرها من المشاكل المتعلقة بواردات المياه الوفيرة(أمطار العواصف والأعاصير) فيمكن استعمال المنحنى لقراءة المجموع السنوي للمطر الذي يفوق احتمالات 10، أو 25 أو 50 في المائة من السنوات أو غيرها من الاحتمالات.

مثال (2):- يتطلب التخطيط لأغراض الإمداد المائي أو التنمية الزراعية معرفة احتمالات سقوط كميات من المطر مساوية لقيمة حرجة أو أقل منها. والجدول (01) التالي يعرض المجاميع السنوية للمطر بمطار طرابلس للفترة 1950- 1979، والمطلوب ما يلي:

1- استخدام ورقة رسم حسابية على المحور السيني واحتمالية التوزيع على المحـور الصـادي لتوقيع قيم التوزيع التكراري التراكمي ثم رسم أفضل خط انحدار يمر بين النقاط بواسطة العين والمسطرة. وهذا الخط سوف يمثل المنحنى التكراري التراكمي للتوزيع ألاعتدالي الذي ينطبق على البيانات.

2- من التوزيع قدر القيم التالية:

أ- المتوسط السنوي للمطر.

ب- الانحراف المعياري.

ج- احتمال سقوط مطر أكبر من 400 ملم.

د- احتمال سقوط أقل من 200 ملم.

جدول (10) المجموع السنوي للمطر بمطار طرابلس خلال الفترة 1950– 1979

265.1	240.8	293.4	228.9	345.8	180.4	245.0	326.5	260.2	379.6
196.0	198.7	356.2	261.2	252.5	471.5	367.9	320.6	285.5	189.6
356.7	357.8	224.4	501.7	304.4	325.2	365.6	266.8	217.0	164.5

الحل:

1- نكون جدول جديد(11) نرتب فيه المفردات تصاعديا ثم نحسب نسب تكرارها F ، ثم نوقع قيم المطر ونسب تكرارها على الشكل البياني (7) الذي يبين التوزيع التكراري للمطر السنوي بمطار طرابلس والذي نستخدمه في الإجابة على بقية الأسئلة.

أ- من الشكل (7) نقدر قيمة المتوسط الحسابي بالتحرك يمينا من النسبة 50 إلى أن نقطع خط الانحدار ثم إلى أسفل نحو قيم المطر فيتبين أن قيمة المتوسط = 290 ملم.

ب- وبنفس الطريقة نتحرك من النسبة 16 يمينا إلى أن نقطع خط الانحدار ثم إلى أسفل فيتبين أن كمية المطر المتوقعة لهذا النسبة = 230 ملم. أما قيمة الانحراف المعياري فهي = 290-230=60 ملم.

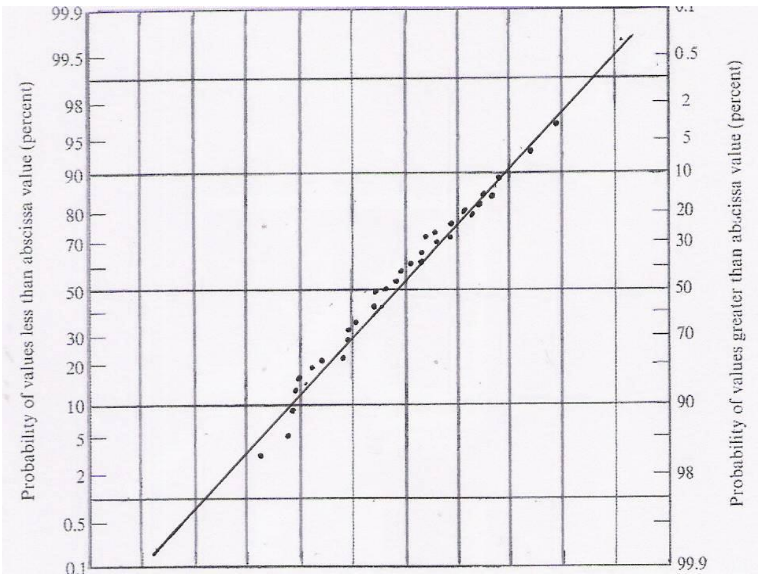
ج- نحصل على احتمال سقوط أكبر من 400 ملم بالصعود من هذا الرقم رأسيا إلى أن نقطع خط الانحدار ثم إلى اليمين فيتبين أن احتمال سقوط كمية أكبر من 400 فهو 10%.

د- احتمال سقوط كمية أصغر من 200 ملم فهو 12%.

جدول(11) حساب تكرار المطر(سنوات) بمنطقة مطار طرابلس للفترة 1950- 1980م.

المطر ملم	الرتبة m	التكرار F	المطر ملم	الرتبة m	التكرار F	المطر ملم	الرتبة m	التكرار F
164.5	1	3.22	252.5	11	35.4	325.2	21	67.7
180.4	2	6.45	260.2	12	38.8	326.5	22	70.9
189.6	3	9.67	261.2	13	41.9	345.8	23	74.1
196.0	4	12.9	265.1	14	45.2	356.2	24	77.4
198.7	5	16.2	265.5.	15	48.4	356.7	25	80.6
217.0	6	19.35	266.8	16	51.2	357.8	26	83.8
224.4	7	22.8	285.5	17	54.8	367.9	27	87.0
228.9	8	25.8	293.4	18	58.0	379.6	28	90.5
240.8	9	29	304.4	19	61.3	471.5	29	93.5
245.0	10	32.2	320.6	20	64.5	501.7	30	96.7

شكل (7) التوزيع التكراري التراكمي للمطر السنوي بمطار طرابلس.



100 200 300 400 500

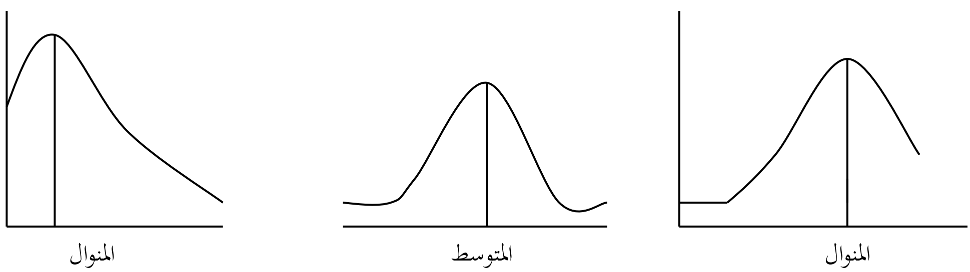
المطر السنوي ملم

الالتواء Skewness:

في كثير من الأحيان نجد أن التوزيع التكراري لمشاهدات العينة لا ينطبق تمام الانطباق على التوزيع ألاعتدالي النموذجي. وقد يكون هذا الانحراف قليلا ليس له دلالة إحصائية فيمكن تجاهله ويعتبر التوزيع في هذه الحالة اعتداليا، لكن في بعض الأحيان يكون الانحراف كبيرا لدرجة لا يستطيع الباحث معها افتراض التوزيع ألاعتدالي في مشاهدات العينة التي يحصل عليها. ففي التوزيعات غير ألاعتدالية قد تلتوي المنحنيات بقدر يؤدي إلى عدم تساوي قيم المتوسط والوسيط والمنوال.

فالتوزيع يوصف بأنه موجب الالتواء عندما تتركز التكرارات عند القيم الصغرى (حيث الغالبية، على سبيل المثال، أغبياء أو فقراء) فيصعد المنحنى بسرعة في البداية ثم يهبط ببطء وهنا يكون المتوسط الحسابي أكبر من الوسيط والمنوال متأثرا بالقيم الكبيرة المتطرفة. ويكون سالب الالتواء عندما تتركز التكرارات في القيم الكبرى(حيث الغالبية أذكياء أو أغنياء) فيصعد المنحنى ببطء ثم يهبط بسرعة ويكون المنوال أكبر من الوسيط والمتوسط .

شكل (8) مقارنة بين التوزيع الاعتدالي والتوزيعات الملتوية.



منحنى سالب الالتواء تركز التكرارات عند القيم الكبيرة

منحنى اعتدالي تتساوى به قيم المتوسط والمنوال والوسيط

منحنى موجب الالتواء تركز التكرارات عند القيم الصغيرة

ولمعرفة نوع الالتواء إما نرسم التوزيع التكراري على ورقة رسم بيانيه كما في الأشكال السابقة أو نستخدم المعادلة التالية :

$$Sk = \frac{\overline{X} - Mo}{S}$$

حيث أن: Sk = (كسر يتراوح بين + 1 و — 1).

= 0 المنحنى اعتدالي لتساوي قيمتي المتوسط والمنوال.

وعندما تكون أكبر من 0 فالمنحنى موجب الالتواء، أما عندما تكون أقل من 0 فالمنحنى سالب الالتواء.

Mo = المنوال (المفردة الأكثر تكرارا في العينة).

S = الانحراف المعياري.

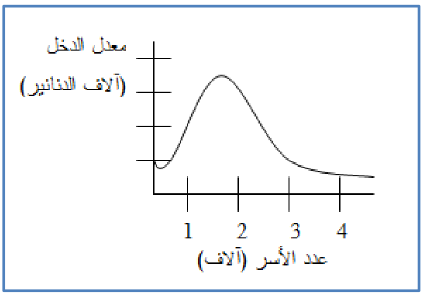
إذن، Sk يوضح مدى اعتدال التوزيع، فحيثما كان المتوسط أكبر من المنوال أصبح الالتواء موجبا وفيه يظهر ذيل المنحنى ملتويا نحو اليمين. أما إذا كان المنوال أكبر من المتوسط تكون Sk أقل من 0 والالتواء سالب، وعندما تتساوى قيم المتوسط والمنوال تكون قيمة Sk مساوية 0 والتوزيع اعتدالي.

والتوزيعات الموجبة الالتواء تتصف بأن مفرداتها الواقعة فوق المتوسط أكثر من المفردات الواقعة تحته . أما في حالة التوزيعات المتماثلة فان أعداد المفردات الواقعة فوق المتوسط تساوي المفردات الواقعة تحته.

ومن التوزيعات الموجبة الالتواء نجد تكرارات الدخل ألسنوي للأسرة، فأغلب الأسر تتركز في المعدلات الصغيرة كما في الشكل التالي:

شكل (9) التوزيع التكراري لفئات الدخل

بآلاف الدنانير في السنة مع عدد الأسر (ألف نسمة).



أمثلة:

1- الجدول التالي يبين المدة بالدقائق التي يستغرقها الطالب في رحلته من البيت إلى كلية الآداب غريان. المطلوب عمل جدول تكراري ثم رسم مضلع تكراري لكي يتبين شكل التوزيع.

جدول(12) المدة بالدقائق التي يستغرقها الطالب في رحلته من البيت إلى الكلية.

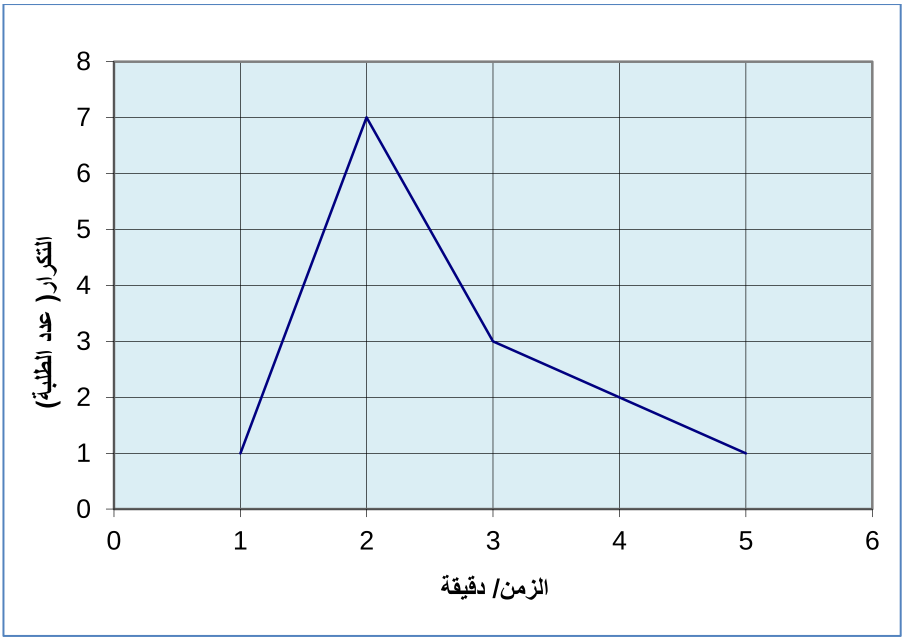
25	45	40	20	20	15	15
15	40	20	15	5	30	25

جدول (13) التوزيع التكراري للزمن الذي يستغرقه الطالب من البيت إلى الكلية.

الفئة	التكرار
1- 10	1
11- 20	7
21- 30	3
31- 40	2
41- 50	1

ومن هذا التوزيع يتبين أن أغلب التكرارات متركزة في القيم الصغرى والذيل ملتوي نحو اليمين لذلك فالتوزيع موجب الالتواء.

شكل (10) التوزيع التكراري للزمن من البيت للكلية



2- استخدم بيانات مدة الرحلة من البيت إلى الكلية في حساب معامل الإلتواء Sk.

الحل: لحساب عناصر المعادلة نكون جدول كما يلي:

جدول (14) المدة بالدقائق التي يستغرقها الطالب في رحلته من البيت الى الكلية.

X	X^2	X	X^2
15	225	25	625
40	1600	45	2025
20	400	40	1600
15	225	20	400
5	25	20	400
30	900	15	225
25	625	15	225

$\sum X$=320

$\sum X^2$=9500

N=14

$$\overline{X} = \frac{330}{14} = 23.57$$

$$S = \sqrt{\frac{9500 - \frac{330^2}{14}}{14-1}} = 11.5$$

ولحساب المنوال نرتب المفردات تصاعديا ثم نختار القيمة الأكثر تكرارا في العينة:

45	40	40	30	25	25	20	20	20	15	15	15	15	5

المفردة الأكثر تكرارا من غيرها هي 15 لذلك فهي المنوال.

والخطوة الأخيرة تتم بالتعويض في المعادلة لحساب معامل الالتواء:

$$Sk = \frac{23.57-15}{11.5} = 0.74$$

الاستنتاج: حيت أن Sk أكبر من 0 فالتوزيع موجب الالتواء.

تمرينات:

1- البيانات التالية لأعلى درجة حرارة عظمى سجلت خلال السنة بمحطة أرصاد زواره للفترة 1950-1982.المطلوب حساب Sk لمعرفة نوع التوزيع .

جدول(15) أعلى درجة حرارة عظمى سجلت خلال السنة بزوارة للفترة 1950-1982.

48.3	45.0	42.8	46.1	42.8	37.2
41.2	48.0	53.0	45.0	54.6	47.2
43.4	46.5	45.0	42.0	44.6	46.2
45.6	43.6	45.0	44.0	44.6	44.3
45.0	39.8	47.5	47.6	43.2	45.0
			45.5	42.5	44.5

2- الجدول التالي يعرض أكبر كميات الأمطار السنوية التي سجلت خلال 24 ساعة في شحات للفترة 1945-1978، والمطلوب حساب معامل Sk لمعرفة نوع التوزيع.

جدول(16) أكبر كميات الأمطار المسجلة خلال 24 ساعة في شحات

103	40	77	34	40	32	35	90	38	49
42	39	63	48	41	34	19	63	40	29
48	35	65	33	70	45	18	49	43	31
						43	81	41	35

أسئلة المراجعة:

1- ما المقصود بالتوزيع التكراري الاعتدالي وما هي خواصه؟

2- ما المقصود بالمنحنى التكراري التراكمي وفيما يستعمل.

3- بين متى يحيد التوزيع التكراري عن الاعتدالي وكيف يمكن التحقق من ذلك؟.

4- متوسط مطر شهر مارس بمحطة أرصاد أحد المراصد يبلغ 9.22 سم بانحراف معياري 2.83 سم، افرض أن كميات المطر تأخذ التوزيع الاعتدالي، أوجد أن يحصل شهر مارس القادم على:

أ- بين 5 و 7 سم.

ب- أقل من 1.84 سم .

ج- أكثر من 13.8 سم .

5- شركة تدفع لعمالها راتبا شهريا بمتوسط 5.25 دينار في الساعة وانحراف معياري 0.60 دينار ، فإذا كانت الرواتب موزعة اعتداليا أوجد نسبة العمال الذين يحصلون على راتب يتراوح بين 4.77 و 5.70 دينار .

6- الجدول التالي يبين مراكز الفئات والتكرارات المتعلقة بمساحات أحواض الصرف النهرية من الدرجة الثانية، اعرض البيانات بأسلوب الأعمدة ثم احسب معامل SK .

جدول (17) أحواض الصرف النهري وتكراراتها.

مركز الفئة قدم^2	100	200	300	400	500	600	700	800	900	1000	1100
التكرار	6	14	7	6	5	3	2	0	1	0	1

الفصل الثالث

الاختبارات المعلمية

أولاً: اختبار المتوسطات:

إن الهدف الأساسي من هذا الفصل هو وضع قواعد على ضوئها يمكن رفض أو قبول جملة افتراضية تخص معالم المجتمع. فعلى سبيل المثال، في أحد الأيام وقبل أن يذهب زيد إلى العمل يتفحص حالة الجو من النافذة فيجد السحب تغطي قبة السماء، فيقول في قرار نفسه: السماء قد تمطر اليوم. إن الذي فعله زيد هو أنه افترض أن السماء قد تمطر أو لا تمطر. فهو ألآن حائر أمام أمرين، إما أن يأخذ المظلة معه أو لا يأخذها وقراره يقوده إلى أربعة نتائج هي:

1- يأخذها وتمطر، والنتيجة هي أنه اتخذ القرار الصحيح.
2- يأخذها ولا تمطر، .... القرار خاطئ.
3- لا يأخذها ولا تمطر، .... القرار صحيح .
4- لا يأخذها وتمطر، ....القرار خاطئ .

إذن، القرارين الأول والثالث صحيحين، فهو يريد المظلة معه إذا السماء أمطرت، ولا يريدها إذا لم تمطر. أما القرارين الثاني والرابع فهما خاطئين، لأن المظلة سوف تسبب له المتاعب بدون فائدة إذا أخذها ولم تمطر، ولا يستفيد منها إذا لم يأخذها وتمطر. ولكن المشكلة هي أنه لا يعرف بالضبط ما إذا كانت ستمطر أم لا، وقراره يسبق الأحداث، وهو مبني على افتراض يمكن أن يكون صحيحا أو خطئا .

إذن، ما هي الفرضية ؟. إنها جملة أو ادعاء بخصوص طبيعة المجتمع . إنها المادة الأولية لأي بحث تجريبي. وفي علم الإحصاء عادة ما تكون الفرضية بخصوص معلم غير معروف يخص المجتمع المدروس . تأمل الحالات التالية:

1- عند رمي قطعة نقود فضية هناك احتمال قدره 50% بأن تستقر على الفارس أو على النجمة . إذن، احتمالية الفارس = 50 في المائة :

$P=0.50$

2- مصنع الإطارات يدعي بأنه يصنع إطارات تسير في المتوسط 40000 كم بدون إعطاب. افرض أن μ هي متوسط عدد الكيلومترات، إذن، الفرضية في هذه الحالة تصاغ كما يلي:

$\mu > 40000$

3- باحث جغرافي يدعي بأن طلبة الجامعة من أصل حضري أكتر تفوقا من الطلبة الريفيين. افرض أن المقياس المستعمل لإجراء الاختبار هو نسبة الخريجين من المدارس الثانوية، فإذا رمزنا لنسبة طلبة المدن Pu ولنسبة طلبة الأرياف Pr إذن الفرضية تصاغ كما يلي:

$Pu > Pr$

والطريقة المثلى التي يمكن بها اختبار الفرضية تتم بدراسة جميع أفراد المجتمع، ونظرا لأن هذا غير ممكن كما سبق وأن أسلفنا عليه نختار عينة نجري عليها الاختبار من أجل الوصول إلى استنتاجات تؤيد الفرضية أو ترفضها.

الفرضية العدمية (الصفرية): Null Hypothesis

وهي الفرضية المختبرة من أجل احتمال الرفض وتختصر في الرمز H0 ودائما تحتوي على إشارة التساوي =. ففي اختبار قطعة النقود افترضنا أن احتمالية الفارس 50% لذلك نكتب الفرضية كما يلي:

$H0: P = 0.50$

الفرضية البديلة: Alternative Hypothesis

عادة ما ترفق فرضية العدم بفرضية أخرى بديلة لها تختصر في الرمز HA وهي الفرضية المقبولة في حالة رفض H0 ، ومجالها لا بد وأن يكون خارج مجال فرضية العدم . ففي مسألة قطعة النقود الفضية تصاغ الفرضية البديلة بأحد الأشكال التالية:

HA: P > 0.50

HA: P < 0.50

HA: P # 0.50

أما في حالة طلبة الجامعة من أصول حضرية وريفية ، فتصاغ فرضية العدم والبديلة كما يلي :

H0 : Pu = Pr

HA : Pu > Pr

العدم تقول بأن ليس هناك فرق بين طلبة المدن والأرياف، أما البديلة فتقول بأن طلبة المدن أفضل. لاحظ أن الفرضية البديلة دائما تؤيد التساؤل الوارد في فكر الباحث أما الفرضية العدمية فهي ترفضه ودائما تنص على التساوي وعدم وجود فرق.

إن اختبار الفرضية الإحصائية يعتبر قاعدة تقود إلى قرار القبول أو الرفض لما جاء في الفرضية H0 وذلك بعد التحصل على معالم العينة. وهذه القاعدة الموضوعة قبل إجراء تحليل العينة تسمى قاعدة القرار decision Role وهي التي يبنى عليها القرار بالقبول أو الرفض، فإذا كانت المؤشرات المجمعة من العينة لا تؤيد الادعاء الوارد في فرضية العدم نرفضه ونستنتج بأن H0 خطأ ونقبل بما جاء في الفرضية البديلة HA .

ولمزيد من التوضيح أفرض أن مهندسا يدعي بأنه طور أسلوبا جديدا يطيل عمر المصابيح التي كانت في السابق تعيش في المتوسط 450 ساعة. الفرضيات:

H0: μ = 450

HA: μ > 450

ومن أجل التبسيط سوف نعالج الحالتين التاليتين:

H0: μ = 450

HA: μ = 500

في هذا المثال وضع المتوسط عند HA = 500 ساعة بالتحديد، ولكي نختبر هذا الادعاء سوف نختار عينة وعلى أساس متوسطها سوف نقرر صحة H0 من عدم صحته. افرض أننا اخترنا عينة متكونة من 10 مصابيح وتحصلنا على متوسط عمرها $\overline{X}$ ، طبعا هناك أخطاء في اختيار العينة تجعل متوسطها لا يساوي 450 ولا 500 بالضبط، لكن إذا كان الرقم المتحصل عليه مثلا 460 أو 445 أو 420 أو أقل فإننا سنحكم لصالح قبول H0 ، أما إذا كان هذا

المتوسط يساوي 490 أو 505 أو 520 أو أكثر فإننا سنرفض H0 ونؤيد HA . لكن ماذا إذا كان يساوي 475 أو 476، في هذه الحالة سوف نبقى مترددين في اتخاذ القرار لصالح H0 أو HA لأنه لا يوجد انحياز كامل لأي منهما . ولكن لابد من أن نضع خطا فاصلا يسمى بالقيمة الحرجة critical value والتي يؤخذ على أساسها القرار بالقبول أو الرفض. افرض أننا اخترنا هذا الفاصل ليكون 475، فإذا كان متوسط العينة أكبر من 475 نرفض H0 ، إذن القيم الأكبر من 475 في هذه الحالة سوف تشكل ما يسمى بالإقليم الحرج أو إقليم الرفض . أما إذا كان متوسط العينة 470 أو أقل فسوف نقرر لصالح H0

ونستنتج بأن المهندس لم يأتي بما يؤيد ادعاءه.

اختبار الفرضيات المتعلقة بمجتمع واحد:

اختبار الطرف الأيمن:

تنصب أغلب الاختبارات الإحصائية على المتوسط الحسابي والانحراف المعياري، ولتبسيط الشرح ذعنا في هذه المقدمة نركز على اختبار المتوسط الحسابي. نود في هذا التحليل اختبار الفرضية التي تقول بأن متوسط المجتمع μ لا يختلف اختلافا ذو دلالة إحصائية عن متوسط افتراضي بل يساويه ($H0: \mu = \mu 0$) ضد الفرضية البديلة التي تنص على أن متوسط المجتمع أكبر من المتوسط الافتراضي ($HA: \mu > \mu 0$). وتحسب قيمة الاختبار الإحصائي بالمعادلة التالية:

$$Z = \frac{u - u0}{\frac{\sigma}{\sqrt{n}}}$$

ونظرا لعدم توفر متوسط المجتمع في أغلب الظروف، نعوض عنه بمتوسط العينة ومن ثم تعدل المعادلة الاختبارية إلى الشكل التالي:

$$Z = \frac{\overline{X} - u0}{\frac{\sigma}{\sqrt{n}}} \qquad \text{.......(1)}$$

حيث أن: $\overline{X}$ = متوسط العينة.

$\mu 0$ = متوسط افتراضي عبارة عن رقم محدد مسبقا.

σ = الانحراف المعياري للمجتمع.

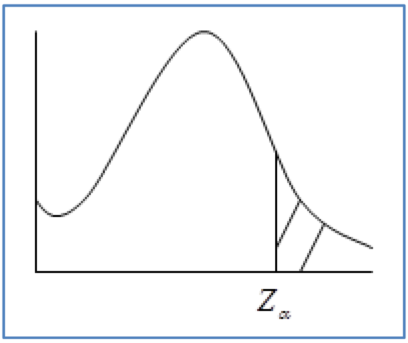
n = حجم العينة.

وبما أن الفرضية البديلة HA تحتوي على إشارة أكبر من فان هذا يحتم علينا تطبيق اختبار الطرف الأيمن. والقاعدة تنص على رفض H0 إذا كانت Z الاختبارية أكبر من Z_α المقدرة من الجدول المعياري رقم (18):

جدول (18) احتمالات التوزيع الزائي Z_α

α	معدل الثقة	الطرف الواحد	الطرفين
0.15	0.85	1.04	1.44
0.10	0.90	1.28	1.65
0.05	0.95	1.65	1.96
0.01	0.99	2.33	2.58

وللتأكد من صحة القرار بالقبول أو الرفض يستحسن رسم المنحنى ألاعتدالي وتحديد موقع Z_α تحته لأجل اتخاذها نقطة للمقارنة مع قيمة Z الاختبارية كما يلي:



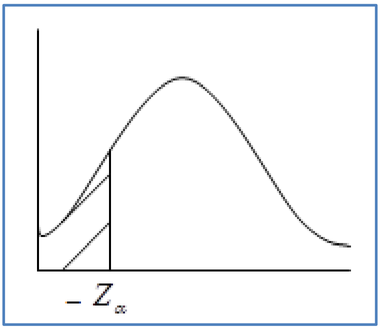
اختبار الطرف الأيسر:

يستخدم هذا الاختبار عندما تحتوي الفرضية البديلة على إشارة أصغر من:

H0: u = u0

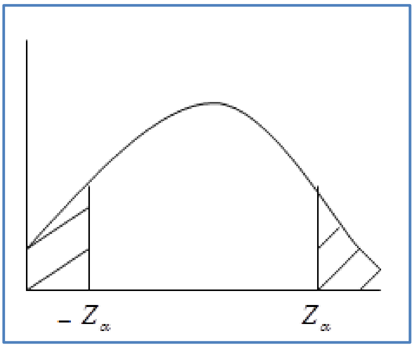
H A: u < u0

والقاعدة تنص على رفض H0 إذا كانت z أصغر من $-Z_\alpha$



اختبار الطرفين :

يستخدم هذا الاختبار عندما تحتوي الفرضية البديلة على إشارة عدم التساوي # ، وفي هذه الحالة فان منطقة الرفض ليست محصورة في طرف واحد كما في الاختبارين السابقين، بل موجودة في الطرفين كما في الشكل التالي :



في اختبار الطرفين تقسم α على طرفي المنحنى بالتساوي، لذلك هنالك منطقتين حرجتين الأولى على اليمين وتساوي $Z_{\frac{\alpha}{2}}$ والثانية على اليسار وتساوي $-Z_{\frac{\alpha}{2}}$. والقاعدة تنص على رفض H0 إذا كانت Z أكبر من $Z_{\frac{\alpha}{2}}$ أو أصغر من $-Z_{\frac{\alpha}{2}}$.

خطوات يجب إتباعها عند اختبار الفرضيات:

1- اقرأ المسألة جيدا ثم أكتب المعطيات المتوفرة أو المطلوب حسابها من العينة مثل:

n =

u0 =

$\sigma =$

$\alpha =$

2- أكتب فرضية العدم التي تنص على أن متوسط المجتمع يساوي متوسطا فرضيا (رقما محددا).

H0 : u = u0

3- أكتب الفرضية البديلة HA والتي تحدد طبيعة الاختبار هل هو للطرف الأيمن أم الأيسر أو للطرفين:

HA : u > u0

HA : u < u0

HA : u # u0

4- اختار عينة عشوائية وأحسب الاختبار الإحصائي:

$$Z = \frac{\overline{X} - u0}{\frac{\sigma}{\sqrt{n}}}$$

5- ارسم المنحنى وبين عليه موقع Z_{α} التي تحصلت عليها من الجدول بناء على معدل α المحدد مسبقا في المسألة .

6- اتخذ القرار بقبول H0 أو رفضها بناء على مقارنة قيمة Z بقيمة $Z\alpha$ المقدرة من الجدول رقم (1) .

اختبار الفرضيات المتعلقة بمتوسط مجتمع واحد معلوم التباين :

إذا كان هناك فرضية خاصة بمتوسط مجتمع معين u فإننا للاستدلال على صحتها نلجأ إلى اختيار عينة من ذلك المجتمع ونحسب متوسطها $\overline{X}$ ثم نقارنه بمتوسط المجتمع المنصوص عليه في الفرضية u0، فإذا كان حجم الفرق بينهما صغيرا أدى ذلك إلى قبول صحة H0 التي تقول بعدم وجود فرق جوهري، وإلا فإننا نرفضها.

أمثلة:

1- افرض أن مجتمعا اعتداليا تباينه يساوي 36 ومتوسطه الحسابي u0 =50، أخذت منه عينة قوامها 16 مفردة متوسطها الحسابي = 48.5 ، أوجد هل متوسط العينة أقل بكثير من متوسط المجتمع، استخدم α = 0.05 .

الحل:

أولاً:- المعطيات:

$$\sigma = \sqrt{36} = 6$$

μ0=50

n=16

$$\overline{X} = 48.5$$

ثانياً:- الفرضيات :

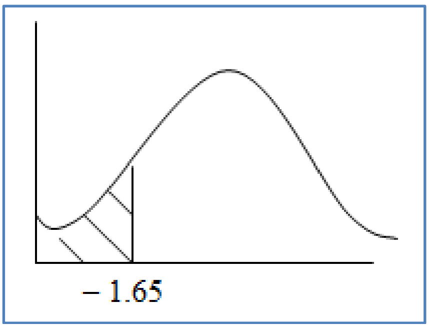
H0 : u = 50

HA : u < 50

ثالثاً:- الاختبار :

$$Z = \frac{48.5 - 50}{\frac{6}{\sqrt{16}}} = -1$$

رابعاً:- تقدير قيمة $Z_{0.05}$ من الجدول المعياري ثم تحديد موقعها تحت المنحنى لأجل اتخاذها أساس للمقارنة مع Z الاختبارية. ومن الجدول يتبين أن قيمة $Z_{0.05}$ = 1.65 ونظرا لأن الاختبار للطرف الأيسر نغير الإشارة فتصبح 1.65 - .



بما أن Z الاختبارية (1 -) ليست بالإقليم الحرج نقبل H0 ونستنتج بأنه ليس هناك فرق جوهري بين متوسط العينة ومتوسط المجتمع .

مثال 2– استخدم بيانات المثال السابق مستبدلا متوسط العينة بالرقم (45) .

الحل: نختصر الخطوات ونحسب الاختبار مباشرة :

$$Z=\frac{45-50}{\frac{6}{\sqrt{16}}}=-3.3$$

من الجدول وجدنا أن قيمة الدرجة المعيارية = 1.65 – وبما أن Z=3.3 – هي بالإقليم الحرج نرفض H0 ونستنتج بأن هناك اختلاف جوهري بين المتوسطين .

مثال 3– متوسط وتباين مبيعات زيد كانا 1000 و 10000 دينار على التوالي، وبعدما أخد دورة في الاقتصاد والتجارة وجد أن مبيعاته خلال 9 أيام عشوائية كما يلي :

1050	950	1100	1300	880	1100	1250	990	1280

استخدم α 0.01 لاختبار هل حصل تحسن في مبيعات زيد نتيجة لدخوله الدورة ؟.

الحل:

أولاً:– المعطيات:

$\mu0=1000$

$\sigma=\sqrt{10000}=100$

n =9

α = 0.01

نحسب متوسط العينة وذلك بقسمة مجموع العينة 9900 على عدد مفرداتها 9 =1100.

ثانياً:– الفرضيات:

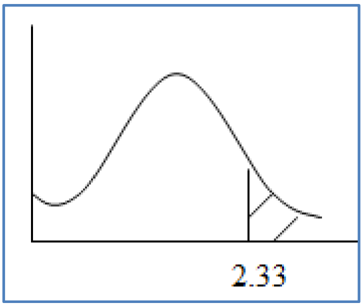
H0: u = 1000

HA : u > 1000

ثالثاً:– الاختبار:

$$Z=\frac{1100-1000}{\frac{100}{\sqrt{9}}}=3.0$$

رابعاً:– قيمة $Z_{0.01} = 2.33$ ، نحدد موقعها تحت المنحنى لأجل تحديد المنطقة الحرجة والمقارنة والاستنتاج :



وبما أن Z واقعة بالإقليم الحرج نرفض H0 ونستنتج بأن الدورة أدت إلى تحسن في مبيعات زيد .

مثال 4– متوسط أوزان أكياس السكر المملوءة بواسطة آلة يبلغ 5 كج بتباين 0.15 كج . ومن أجل اختبار درجة كفاءة الآلة، أخذت عينة عشوائية من 16 كيسا وتبين أن متوسط أوزانها 5.1 كج . استخدم α 0.05 لاختبار هل يوجد ما يدل على أن درجة كفاءة الآلة تختلف كثيرا عن السابق .

الحل :

أولاً:– المعطيات :

μ0=5

n=16

$$\sigma = \sqrt{0.15} = 0.3872$$

$$\overline{X} = 5.1$$

ثانياً:– الفرضيات :

H0 = u = 5

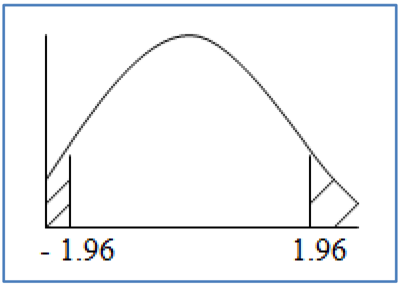
HA = u # 5

ثالثاً:– الاختبار:

$$Z = \frac{5.1 - 5}{\frac{0.3872}{\sqrt{16}}} = 1.03$$

رابعاً:– تحديد المنطقة الحرجة تحت المنحنى لأجل المقارنة والاستنتاج:

قيمة $Z_{0.05}$ لاختبار الطرفين = 1.96 أو– 1.96 وبما أن القيمة الاختبارية = 1.03 ليست بالإقليم الحرج نقبل H0 ونستنتج أن الآلة معدلة فهي تملأ الأكياس بالمعدل المصممة من أجله .



تمرينات :

1– متوسط الكسر في قنينات أحد شركات صناعة الزجاج يبلغ 15 قطعة وبتباين 4 قطعة في الصندوق أستعمل أسلوب جديد في التعبئة يعتقد أنه يقلل من معدل الكسر، وللتأكد من سلامة هذا الأسلوب درست عينة متكونة من 40 صندوق فتبين أن متوسط الكسر بها يساوي12 قطعة في الصندوق، استخدم α 0.05 لاختبار هل حصل تحسن كبير في معدل الكسر يبرر الاستمرار في استعمال هذا الأسلوب المكلف ماليا.

2– مصنع للأجهزة الرياضية يدعي بأنه طور أسلاك جديدة لصيد السمك متوسط قوة القطع لديها 8 كج بانحراف معياري 0.5 كج، اختيرت عينة عشوائية من 60 سلكا تبين أن متوسط قوتها 8.7 كج، استخدم α 0.05 لاختبار ادعاء المصنع.

3– معدل الوقت الذي كان يستغرقه الطالب للتسجيل بالجامعة الأمريكية 50 دقيقة، ولتقصير هذه المدة أستخدم الحاسوب. وللتأكد من مدى كفاءة هذا الأسلوب اختيرت عينة عشوائية متكونة من 15 طالب تبين أن متوسط المدة التي استغرقوها في التسجيل 34 دقيقة. استخدم α 0.01 لاختبار هل وقت التسجيل قل جذريا عما كان عليه قبل استعمال الحاسوب.

اختبار الفرضيات المتعلقة بمتوسط مجتمع واحد تباينه غير معلوم والعينة صغيرة:

عند اختبار الفرضيات المتعلقة بمتوسط المجتمع μ في الغالب لا نعرف تباين المجتمع σ^2 ولكن توجد الرغبة في اختبار الفرضية التي تنص على أن متوسط المجتمع μ لا يختلف عن قيمة محددة $\mu 0$، أي :

$$H0 : \mu = \mu 0$$

وفي حالة وجود تباين المجتمع σ^2 نستخدم الاختبار الزائي، أما عندما لا يتوفر تباين المجتمع نستعيض عنه بتباين العينة S^2 ونستخدم الاختبار التائي بدلا من الزائي كما يلي:

$$T = \frac{\overline{X} - \mu 0}{\frac{S}{\sqrt{n}}} \qquad (2)$$

<u>إذن، الفرق الوحيد بين الاختبارين Z و T هو كون الأول يعتمد في حسابه على تباين المجتمع، بينما يعتمد الثاني على تباين العينة .</u>

وبعد حساب T الإختبارية نقارنها بقيمة T_α المقدرة من الجدول المعياري بدرجات الحرية (n-1)=df .

قواعد الاختبار:- ارفض H0 إذا كانت T :

1- أكبر من (على يمين) T_α في اختبار الطرف الأيمن.

2- أصغر من (على يسار) $-T_\alpha$ في اختبار الطرف الأيسر .

3- أكبر من (على يمين) $T_{\frac{\alpha}{2}}$ أو أصغر من (على يسار) $-T_{\frac{\alpha}{2}}$ في اختبار الطرفين .

أمثلة :

1- أحد مصانع السيارات يدعي بأن سياراته الجديدة تسير في المتوسط أكثر من 10 كم لكل لتر من البنزين ، ولاختبار هذا الادعاء أجريت تجارب على 10 سيارات زود كل منها بلتر من البنزين فكانت النتائج كما يلي :

11	14	8	8	11	9	9	12	8	11

استخدم α 0.05 لاختبار مدى صحة هذا الادعاء .

الحل:

أولاً:- المعطيات :

$\mu 0 = 10$

$n = 10$

$\alpha = 0.05$

$\Sigma x = 101$

$\Sigma X^2 = 1057$

بما أن تباين المجتمع غير موجود نستعيض عنه بتباين العينة، كما أن متوسط العينة غير محسوب فيجب تقديره أيضا :

$$\overline{X} = \frac{101}{10} = 10.1$$

$$S = \sqrt{\frac{1057 - \frac{101^2}{10}}{10 - 1}} = 2.02$$

ثانياً:- الفرضيات :

$H0 : \mu = 10$

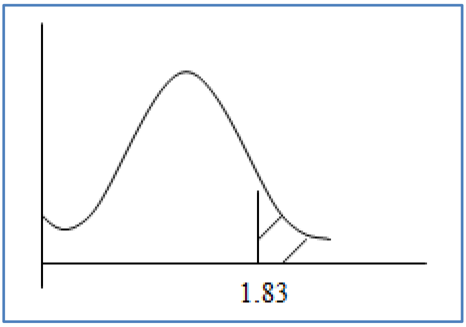
$HA : \mu > 10$

ثالثاً:- الاختبار :

$$T = \frac{10.1 - 10}{\frac{2.02}{\sqrt{10}}} = 0.16$$

رابعاً:- تحديد الإقليم الحرج تحت المنحنى والمقارنة والاستنتاج :

قيمة T الجدولية المقدرة لمعدل α 0.05 و 9 درجات حرية = 1.83 . ألآن نبين موقعها تحت المنحنى لأجل المقارنة والاستنتاج :



بما أن T الاختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأن التجربة لم تؤكد ادعاء المصنع .

مثال 2– صممت آلة لملأ أكياس الدقيق بوزن 5 كج وبعد سنة أخذت عينة عشوائية متكونة من 6 أكياس أوزانها كما يلي :

5.3	4.8	5.2	4.8	5.2	5.3

عند مستوى ثقة 0.95 أوجد هل الآلة مازالت معدلة أم لا .

الحل:

أولاً:– المعطيات :

n = 6

u0 = 5

$$S = \sqrt{0.056} = 0.2366$$

$$\overline{X} = \frac{30.6}{6} = 5.1$$

ثانياً:– الفرضيات :

H0 : u = 5

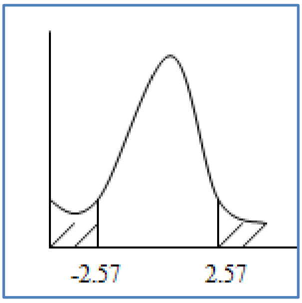
HA : u # 5

ثالثاً:– الاختبار :

$$T = \frac{5.1-5}{\frac{0.2366}{\sqrt{6}}} = 1.04$$

رابعاً:– تحديد الإقليم الحرج تحت المنحنى والمقارنة والاستنتاج :

قيمة $T_{0.05}$ و 5 df لاختبار الطرفين = 2.57



وحيت أن T الاختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأن الآلة مازالت معدلة .

مثال 3– صيدلي يدعي بأنه طور دواء جديدا لتخفيف الوزن ، وللتحقق من هذا الادعاء أخذت عينة عشوائية متكونة من 5 أشخاص استعملوا هذا الدواء ، والبيانات التالية هي لأوزانهم بعد أخذهم للدواء وقبله . عند مستوى ثقة 0.95 هل هناك ما يؤيد ادعاء الصيدلي .

رقم الشخص	1	2	3	4	5
الوزن بعد	75	69	64	65	79
الوزن قبل	80	70	65	62	83
فرق الوزن	−5	−1	−1	3	−4

الحل :

أولاً:– المعطيات :

n=5

u0=0

$$\overline{X} = \frac{-8}{5} = -1.6$$

$$S = \sqrt{\frac{52 - \frac{(-8)^2}{5}}{5-1}} = 3.13$$

ثانياً:– الفرضيات :

H0 : u = 0

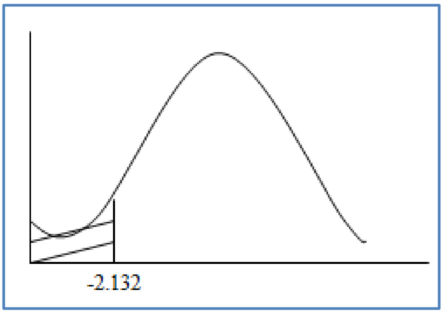
HA : u < 0

ثالثاً:- الاختبار :

$$T = \frac{-1.6 - 0}{\frac{3.13}{\sqrt{5}}} = -1.14$$

رابعاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج :

القيمة الجدولية المقدرة عند ألفا و 4 درجات حرية = 2.132 - ، وبما أن قيمة T الاختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأنه ليس هناك ما يؤيد ادعاء الصيدلي، فالدواء وان كان في مجمله يخفض الوزن، إلا أن هذا التخفيض ليس ذو أهمية إحصائية .



تمرينات :

1- متوسط أوزان نوع من العجول في نهاية الشهر الثالث من العمر يبلغ 65 كج ،عزلت عينة من 12 عجل عند الولادة وأعطيت غذاء خاصا إلى أن وصل عمرها أربعة أشهر وتبين أن أوزانها كما يلي :

67 65 69 66 63 65 76 70 69 64 68 65

عند مستوى دلالة 0.95 أوجد هل أدى هذا الغذاء الخاص إلى تحسن في أوزان العجول .

2- مصنع للسيارات يدعي بأن سياراته الجديدة سوف تسير لمسافة تزيد في المتوسط عن 20000 كم بدون حدوث أعطال ، وللتأكد من ذلك أجريت تجارب على 4 سيارات وكانت نتائجها كالتالي :

21000 20000 22000 19000 كم

أختبر عند مستوى دلالة 0.95 صحة هذا الادعاء من عدمه .

3-لاختبار تأثير نوع من العقاقير الطبية على معدل ضغط الدم، قام باحث بقياس الضغط لدى عينة من 12 رجل قبل تعاطي الدواء وبعده . والبيانات مرتبة في الجدول التالي، والمطلوب اختبار هل هذا الدواء يزيد من الضغط، استخدم مستوى دلالة 0.95 .

ملاحظة: إذا لم يكن هناك تأثير للدواء ففارق الضغط = 0

الرقم	1	2	3	4	5	6	7	8	9	10	11	12
الضغط بعد	128	127	118	135	131	141	135	129	131	132	137	132
الضغط قبل	120	118	126	127	130	140	128	126	124	140	135	130
الفرق	8	9	-8	8	1	1	7	3	7	-8	2	2

4- تدعي إحدى شركات التبغ بأن متوسط محتوى سجائرها من النيكوتين أقل من 14 ملغ، وللتحقق من ذلك قام مهندس الرقابة الصناعية بتحليل محتويات 5 سجاير فوجد أن كمية النيكوتين فيها كما يلي :

26 28 22 23 29 ملغ

هل هناك ما يؤيد ادعاء الشركة، استخدم α 0.05 .

5- مصنع للمصابيح الكهربائية يدعي بأن متوسط عمر المصابيح التي يصنعها 500 ساعة، أخذت عينة من 25 مصباحا وتبين أن متوسطها 518 ساعة وانحرافها المعياري 40 ساعة. هل هناك اختلاف ذو دلالة إحصائية بين متوسط العينة وبين ما ادعاه المصنع α =0.05 .

ثانياً: اختبار تباين المجتمع σ^2 :

إن معرفة معامل التباين مهم جدا في كثير من المسائل التطبيقية، مثل مراقبة الجودة في الإنتاج، حيت أن حجم التباين يوضح فيما إذا كانت الآلة معدلة حسب المواصفات أم لا .
أفرض أن آلة تملأ الأكياس بوزن 5 كيلوجرامات بالضبط حسب المواصفات القياسية، ولكن تعتبر غير معدلة إذا زاد الانحراف المعياري أو قل عن 1 كيلوجرام. إذن، من السهل ادراك أهمية الانحراف المعياري وكيف أنه يعتبر الهدف الرئيسي للتحليل في الكثير من البحوث التطبيقية .

في هذا الاختبار نحن نرغب في اختبار الفرضية H0 التي تنص على أن تباين المجتمع لا يزيد عن قيمة محددة نرمز لها σ_0^2 بل مساويا لها، ضد الفرضيات البديلة :

$$H0 : \sigma^2 = \sigma_0^2$$

$$HA : \sigma^2 > \sigma_0^2$$

$HA: \sigma^2 < \sigma_0^2$

$HA: \sigma^2 \# \sigma_0^2$

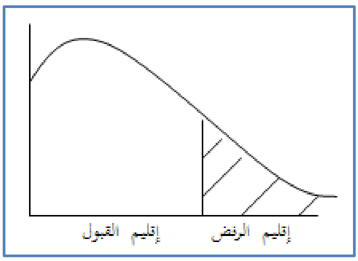
وفي أغلب الأحيان لا يتوفر لدينا تباين المجتمع لذلك نستعيض عنه بتباين العينة S^2 ، والاختبار الإحصائي المناسب لهذه الحالة هو مربع كآي :

$$\chi^2 = \frac{(n-1)}{\sigma_0^2} S^2 \quad \text{...(3)}$$

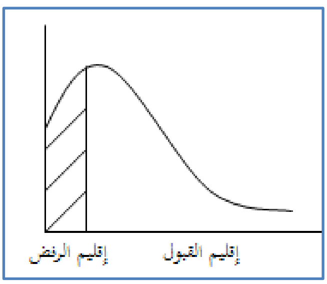
وكما عرفنا في الاختبارات السابقة تقارن قيمة مربع كآي الاختبارية بالقيمة المعيارية المقدرة من الجدول باستخدام α ودرجات الحرية (df=n-1).

قواعد الاختبار :

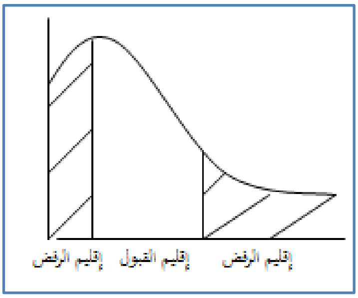
1- **لاختبار الطرف الأيمن** : أوجد قيمة مربع كآي χ^2 عند α و (df=n-1) من الجدول ثم حدد موقعها تحت المنحنى وارفض H0 إذا كانت قيمة مربع كآي الاختبارية أكبر من القيمة الجدولية، أي (بالإقليم الحرج).



2- **لاختبار الطرف الأيسر** أوجد قيمة مربع كآي المقابلة لدرجات الحرية و القيمة ($1-\alpha$)، ثم ارفض H0 إذا كانت القيمة الاختبارية أصغر من (على يسار) القيمة الحرجة .



3- **لاختبار الطرفين** تقسم α إلى قسمين لذلك تحدد قيمة مربع كآي للطرف الأيمن بالقيمة المقابلة لدرجات الحرية df و $\frac{\alpha}{2}$ أما للطرف الأيسر فتحدد بالقيمة المقابلة لدرجات الحرية و $\left(1-\frac{\alpha}{2}\right)$. أرسم المنحنى وبين عليه الموقعين الحرجين عند الطرفين الأيمن والأيسر ثم أرفض H0 إذا وقعت χ^2 في أي منهما، وبالعكس أقبل H0 إذا وقعت بإقليم القبول .



1- فيما يلي أوزان 6 فئران تجارب مختارة عشوائيا، أوجد هل تباين أوزان الفئران أكبر من 63 جرام، استخدم مستوى ثقة 0.95 .

225	340	350	200	320	300

الحل : بما أن الحديث يدور حول التباين فالاختبار المناسب هو مربع كآي :

أولاً:- المعطيات المتوفرة والمحسوبة من العينة :

$n = 6$

$\sum X^2 = 521125$

$\sum X = 1735$

$\sigma_0^2 = 63$

$$S^2 = \frac{521125 - 501704.17}{6-1} = 3884.16$$

ثانياً:- الفرضيات:

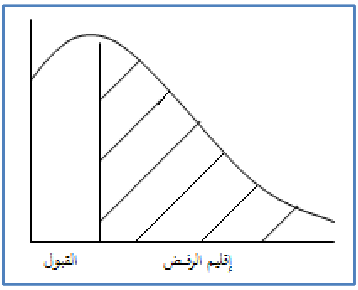
$H0: \sigma^2 = 63$

$HA: \sigma^2 > 63$

ثالثاً:- الاختبار :

$$\chi^2 = \frac{(6-1)}{63} 3884.16 = 308.3$$

رابعاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج : القيمة الجدولية لكآي تربيع = 11.07



وحيت أن كآي تربيع الإختبارية أكبر من الجدولية (واقعة بالإقليم الحرج) نرفض H0 ونستنتج بأننا 95% متأكدين بأن تباين الفئران أكبر من 63 جرام .

مثال 2- مصنع الحديد والصلب يدعي بأن تباين الأنابيب التي ينتجها أقل من 0.1 سم، وللتأكد من ذلك، أخذت عينة من 11 أنبوب، تبين أن انحرافها المعياري = 0.07 سم .استخدم α 0.05 لاختبار مدى صحة هذا الادعاء .

الحل :

أولاً:- المعطيات :

$n = 11$

$\sigma^2 = 0.1$

$S^2 = 0.07$

ثانياً:- الفرضيات :

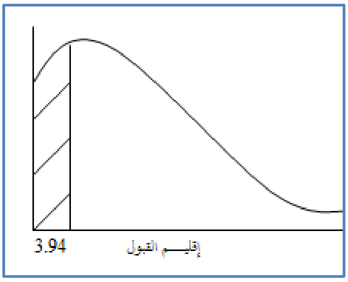
$H0: \sigma^2 = 0.1$

$HA: \sigma^2 < 0.1$

ثالثاً:- الاختبار :

$$\chi^2 = \frac{(11-1)}{0.1} 0.07^2 = 0.49$$

ولتحديد قيمة كآي تربيع المعيارية ننتقل إلى الجدول بدرجات حرية = 10 ونظرا لأن الاختبار للطرف الأيسر فإننا نبحث عن الدرجة التي تقابل ($1 - \alpha$) أي تحت 0.95 ونجد أنها = 3.94 ، الآن نرسم المنحنى ونبين عليه المنطقة الحرجة الواقعة إلى يسار 3.94



رابعاً:- المقارنة والاستنتاج: حيث أن كآي تربيع ليست واقعة بالإقليم الحرج نقبل H0 ونستنتج بأننا 95% متأكدين بأن أداء المصنع صحيح .

مثال 3- عينة عشوائية متكونة من 25 مصباحا كهربائيا وجد أن الانحراف المعياري لمدة تشغيلها 13 ساعة. هل تباين العينة يختلف جذريا عن تباين المجتمع الذي أخذت منه وهو 100 ساعة، استخدم α 0.05 .

الحل :

أولاً:- المعطيات

n = 25

$S^2 = 13^2 = 169$

$\sigma^2 = 100$

a=0.05

ثانياً:- الفرضيات:

$H0: \sigma^2 = 100$

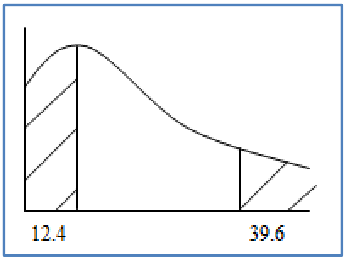
$HA: \sigma^2 \# 100$

ثالثاً:- الاختبار :

$$\chi^2 = \frac{(25-1)}{100} 13^2 = 40.56$$

هناك قيمتين لكآي تربيع الأولى للطرف الأيمن وتحدد بدرجات الحرية و $\frac{\alpha}{2}$ =39.6 والثانية للطرف الأيسر وتحدد بدرجات الحرية و $\left(1-\frac{\alpha}{2}\right)$ = 12.4

رابعاً:- تحديد المنطقة الحرجة والمقارنة والاستنتاج :



بما أن كآي تربيع واقعة بالإقليم الحرج الأيمن نرفض H0 ونستنتج بأن تباين المصابيح يختلف جذريا عن 100 ساعة .

تمرينات :

1- إذا كان تباين درجات امتحان الذكاء المقنن =50 ، وإذا أخذت عينة عشوائية قوامها 30 فردا أعطوا الامتحان المذكور ووجد أن تباين درجاتهم =90 ، هل هناك ما يدل على أن تباين العينة يختلف عن تباين المجتمع المأخوذة منه، استخدم معدل ثقة 0.95.

2- مصنع للنضائد يعلن بأن نضائده تعيش في المتوسط 3 سنوات وبانحراف معياري 1 سنة، اختيرت منها عينة عشوائية ووجد أن أعمارها كما يلي : 1.9 2.4 3.0 3.5 4.2 سنوات . استخدم α 0.05 لاختبار هل الانحراف المعياري ما زال لم يختلف عن 1 سنة .

ثالثاً: اختبار نسبة المجتمع:

الهدف الأساسي من هذا الدرس هو اختبار الفرضيات المتعلقة بالنسبة المئوية لخاصية معينة في العينة. فاختبار النسب مطلوب في كثير من حقول المعرفة ،فالرجل السياسي يهمه معرفة كم هي نسبة المصوتين له في الانتخابات القادمة، وإدارات المصانع يهمها معرفة نسبة المنتجات الفاسدة الخ ..

الفرضيات :

H0 : P = p0

HA : P > P0

: P < P0

: P # P0

حيث أن : P0 نسبة افتراضية، وهي عبارة عن كسر عشري واقع بين 0 و 1

P النسبة المحسوبة من العينة أي نسبة الأفراد الذين تنطبق عليهم الحالة، وتحسب بالمعادلة :

$$P = n / N \qquad (4)$$

n = عدد الأفراد الذين تنطبق عليهم الحالة .

N = حجم العينة .

والاختبار الزائي هو القانون المستخدم للاختبار النسب ويحسب بالمعادلة التالية :

$$Z = \frac{P - P0}{\sqrt{\frac{P0(1-P0)}{N}}} \qquad (5)$$

قواعد الاختبار : ارفض H0 إذا كانت Z أكبر من Z_α (على يمينها) في اختبار الطرف الأيمن، وأصغر من $-Z_\alpha$ (على يسارها) في اختبار الطرف الأيسر. وأكبر من $Z_{\frac{\alpha}{2}}$ أو أصغر من $-Z_{\frac{\alpha}{2}}$ في اختبار الطرفين .

مثال 1– آلة معروفة بأنها تصنع 0.30 من الأنابيب معيبة ، وبعد عملية صيانة وجد أنها تصنع 0.22 بها عيوب، هل صحيح أن نسبة الفساد قلت جوهريا بعد عملية الصيانة، استخدم معدل ثقة 0.95 .

الحل :

أولاً:– المعطيات والفرضيات :

P0=0.30

P=0.22

$\alpha = 0.05$

H0: P=0.30

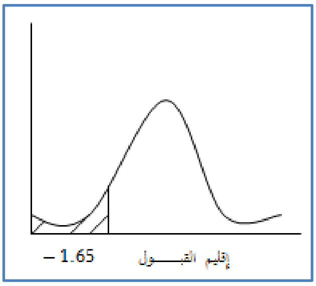
HA : P<0.30

ثانياً:– الاختبار :

$$Z = \frac{0.22 - 0.30}{\sqrt{\frac{0.30(1-0.30)}{100}}} = -1.745$$

ثالثاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج :

من الجدول المعياري يتبين أن قيمة Z_α =1.65 - ، نحدد موقعها تحت المنحنى للمقارنة :



بما أن Z واقعة بالإقليم الحرج نرفض H0 ونستنتج بأننا 95% متأكدين بأن نسبة الفساد قد قلت عن ذي قبل .

مثال2- بلغت نسبة المسافرين للخارج السنة الماضية 20% ، ولمعرفة رأي الناس في السفر تم مقابلة 100 شخص هذه السنة أجاب 15 منهم بأنهم يرغبون في السفر والباقي لا يرغبون، استخدم معدل ثقة 90% لاختبار هل نسبة رغبات الناس قد تغيرت عن السنة الماضية .

الحل:

أولاً:- المعطيات و الفرضيات :

P0=0.20

P=0.15

N=100

$\alpha = 0.10$

H0 : P = 0.20

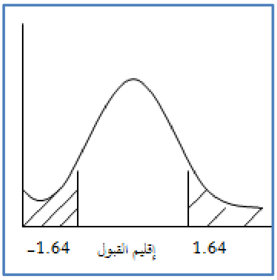
HA : P# 0.20

ثانياً:- الاختبار :

$$Z = \frac{0.15 - 0.20}{\sqrt{\frac{0.20(1-0.20)}{100}}} = -1.25$$

ثالثاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج :

الفرضية البديلة هي لاختبار الطرفين لذلك هنالك قيمتين معياريتين يجب تقديرهما ومن الجدول يتبين أن قيمتي $Z_{0.10}$ = 1.645 و 1.645 - ، وحيث أن Z الاختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأن رغبات الناس لم تتغير جذريا عن السنة الماضية .



تمرينات:

1- بلغ عدد الأشخاص المسافرين للخارج السنة الماضية 200 لكل 1000 من السكان أما عددهم هذه السنة بلغ 60 لكل 200 نسمة، استخدم معدل ثقة 95% لمعرفة هل نسبة المسافرين قد زادت عن السنة الماضية .

2- أحصي مواليد شهر مارس في أحد المدن فكان يساوي 500 من بينهم 210 أنثى والباقي ذكور هل تدل هذه البيانات على أن نسبة الإناث تختلف جذريا عن 50%، استخدم مستوى ثقة 95% .

3- يدعي أنصار التعليم المبرمج أن استعمال الكمبيوتر في تعليم الإحصاء يزيد نسبة النجاح إلى 75% فإذا تعلم 27 طالبا باستعمال هذا البرنامج ونجح منهم 24، فهل هذه النتائج تؤيد صحة الادعاء الذي يروج له أنصار التعليم المبرمج، استخدم α 0.05 .

4- يسود اعتقاد بأن إتباع أسلوب جديد في التعليم يقلل نسبة الفاقد إلى 0.06 ، تم تطبيق هذا الأسلوب على 500 طالب في عدد من المدارس وكان عدد الفاقد 30 طالبا، استخدم معدل ثقة 95% للتأكد هل هناك ما يؤيد هذا الاعتقاد .

5- تدعي شركة للأدوية بأنه على الأقل 80% من الأشخاص الذين يأخذون الدواء الجديد سوف يكتسبون مناعة ضد السعال، اختيرت عينة من 125 شخصا من الذين تناولوا الدواء وتبين أن عدد الذين اكتسبوا المناعة بلغ 90 شخصا، اختبر صحة هذا الادعاء عند مستوى ثقة 95%.

رابعاً: اختبار الفرضيات المتعلقة بمجتمعين إحصائيين:

تناولنا في الصفحات السابقة اختبار الفرضيات المتعلقة بخاصية معينة من خصائص مجتمع واحد مع اختلاف الزمن، أي كيف كانت وكيف أصبحت. أما في هذا الموضوع فسوف نتناول بالشرح الفرضيات المتعلقة بمجتمعين إحصائيين في وقت واحد .

1- اختبار الفرق بين متوسطي مجتمعين تباين كل منهما معلوم :

في هذا الموضوع نرغب في اختبار الفرضية التي تنص على عدم وجود فرق ذو دلالة إحصائية بين متوسطي مجتمعين ضد الفرضية البديلة HA القائلة بوجود فرق ذو دلالة إحصائية بينهما .

$H0 : \mu1 = \mu2$

$HA : \mu1 > \mu2$

$: \mu1 < \mu2$

$: \mu1 \# \mu2$

والاختبار الإحصائي هو :

$$Z = \frac{\mu1 - \mu2}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

وحيث أن متوسطي المجتمعين $\mu1$ و $\mu2$ في الغالب غير متوفرين نستعيض عنهما بمتوسطي العينتين

$\overline{X}$ و $\overline{Y}$ لذلك تعدل المعادلة الاختبارية كما يلي :

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \quad \text{...(6)}$$

قواعد الاختبار: ارفض H0 إذا كانت Z الاختبارية :

1- أكبر من (على يمين) Z_α في اختبار الطرف الأيمن .

2- أصغر من (على يسار) $-Z_\alpha$ في اختبار الطرف الأيسر .

3- أكبر من $Z_{\frac{\alpha}{2}}$ أو أصغر من $-Z_{\frac{\alpha}{2}}$ في اختبار الطرفين .

أمثلة :

مثال1– عينة عشوائية متكونة من 15 رجلا أوروبيا متوسط وزنها 75 كج، بينما بلغ متوسط وزن عينة أمريكية متكونة من 18 رجلا 78 كج، ومن دراسة سابقة وجد أن تباين الأوروبيين 40 كج وتباين الأمريكيين 50 كج . استخدم مستوى ثقة 95% لاختبار هل هناك اختلاف جوهري بين الوزنين .

الحل:

أولاً:– المعطيات والفرضيات :

الأوروبيين	**الأمريكيين**
$m = 15$	$n = 18$
$\overline{X} = 75$	$\overline{Y} = 78$
$\sigma_1^2 = 40$	$\sigma_2^2 = 50$

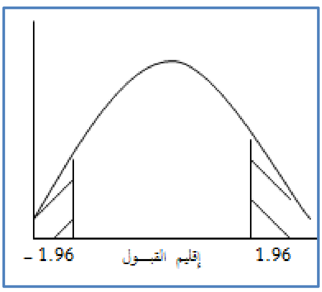
$$H0 : \mu 1 = \mu 2$$

$$HA : \mu 1 \# \mu 2$$

ثانياً:– الاختبار : بما أن تبايني العينتين معلوم فالاختبار المناسب هو Z:

$$Z = \frac{75 - 78}{\sqrt{\left(\frac{40}{15} + \frac{50}{18}\right)}} = -1.28$$

ثالثاً:– تحديد الإقليم الحرج والمقارنة والاستنتاج :



بما أن Z ليست بالإقليم الحرج نقبل H0 ونستنتج بأن ليس هناك اختلاف جدري بين المتوسطين الأوروبي والأمريكي .

مثال 2- قام فلكيان برصد الزمن الذي يستغرقه أحد النجوم ليمر أمام عدسة المرصد، فإذا كان عدد المشاهدات التي قام بها الأول = 12 بمتوسط 1.2 ثانيه لكل مشاهده، وعدد المشاهدات التي قام بها الثاني =8 بمتوسط 1.15 ثانية لكل مشاهدة، ومن الخبرة السابقة معروف بأن تباين مشاهدات الأول والثاني = 0.4 و 0.3 على التوالي، استخدم مستوى دلالة 99% لاختبار هل هناك اختلاف بين مشاهدات الاثنين .

الحل :

أولاً:- المعطيات والفرضيات :

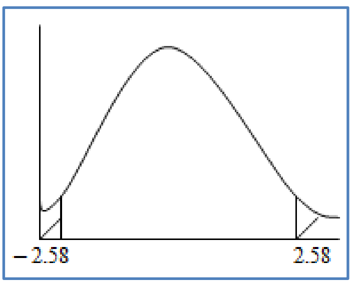
الأول	الثاني
$m = 12$	$n = 8$
$\overline{X} = 1.2$	$\overline{Y} = 1.15$
$\sigma_1^2 = 0.4$	$\sigma_2^2 = 0.3$

$$\alpha = 0.01$$
$$H0: \mu 1 = \mu 2$$
$$HA: \mu 1 \# \mu 2$$

ثانياً:- التباين معلوم مسبقا لذلك نستخدم الاختبار الزائي :

$$Z = \frac{1.2 - 1.15}{\sqrt{\frac{0.4}{12} + \frac{0.3}{8}}} = 0.1878$$

ثالثاً:- تحديد الإقليم الحرج (الرفض) والمقارنة والاستنتاج :



قيمة $Z_{0.01}$ لاختبار الطرفين = 2.58 و - 2.58 وبما أن Z الاختبارية (0.1878) واقعة بإقليم القبول نقبل H0 ونستنتج بأنه ليس هناك فرق معتبر بين مشاهدات الفلكيين .

تمرينات :

1- أجرى اختبار لعينتين الأولى متكونة من 40 طالبا متوسط درجاتها 76 والثانية متكونة من 50 طالبة متوسط درجاتها 81 ، ومن الخبرة السابقة تأكد أن تباين درجات الطلاب = 11 وتباين درجات الطالبات = 9 ، فهل هناك ما يدعو إلى القول بأن الطالبات متفوقات على الطلبة، α =0.05 .

2- استخدم بيانات التمرين السابق لاختبار الفرضية البديلة HA التي تنص على وجود اختلاف كبير بين درجات الطلبة والطالبات، α = 0.01

2- اختبار الفرق بين متوسطي مجتمعين تباينهما غير معلوم :

الاختبار التالي يستخدم في حالة وجود عينتين صغيرتين مختارتين من مجتمعين تباينهما غير معلوم، وكما بيننا في السابق أنه في حالة عدم توفر تباين المجتمع نقدره من العينة ونستخدم اختبار T بدلا من Z .

خطوات الاختبار :

أولاً:- الفرضيات :

$$H0: \mu1 = \mu2$$
$$HA: \mu1 > \mu2$$
$$HA: \mu1 < \mu2$$
$$HA: \mu1 \# \mu2$$

ثانياً:- الاختبار :

$$T = \frac{\mu1 - \mu2}{Sp\sqrt{\frac{1}{m} + \frac{1}{n}}} \quad ......(7)$$

حيث أن Sp = الانحراف المعياري المرجح من العينتين ويقدر بالمعادلة التالية :

$$Sp = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}} \quad ......(8)$$

ثالثاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج :

من الجدول نقدر قيمة T المعيارية عند α ودرجات الحرية df التي تساوي حاصل جمع العينتين مطروح منه 2 أي : (m+n-2) .

قواعد الاختبار: أرفض H0 إذا كانت :

1-T أكبر من (على يمين) T_{α} في اختبار الطرف الأيمن .

2- T أصغر من (على يسار) $-T_{\alpha}$ في اختبار الطرف الأيسر .

3- T أكبر من $T_{\frac{\alpha}{2}}$ أو أصغر من $-T_{\frac{\alpha}{2}}$ في اختبار الطرفين .

أمثلة :

1- يراد المفاضلة بين طريقتين في تعلم اللغة الأجنبية فاختيرت عينتين متكونتين من 12 طالبا لكل منهما، والدرجات التي تحصل عليها أفراد كل عينة هي كما يلي :

رقم الطالب	1	2	3	4	5	6	7	8	9	10	11	12
الطريقة الأولى	26	34	34	30	38	32	31	35	29	29	31	12
الطريقة الثانية	29	24	26	31	32	30	36	29	28	29	28	32

هل هناك اختلاف جوهري يؤدي إلى تفضيل إحدى الطريقتين على الأخرى ، α = 0.05 .

الحل :

أولاً:- المعطيات المتوفرة والمحسوبة والفرضيات :

الأولى الثانيـة

$\overline{X} = 31.75$..................................$\overline{Y} = 29.5$

$m = 12$..$n = 12$

$S_1^2 = 10.2$..................................$S_2^2 = 6.2$

$df = 12 + 12 - 2 = 22$

$\alpha = 0.05$

$HO: u1 = u2$

$HA: \mu1 \# \mu2$

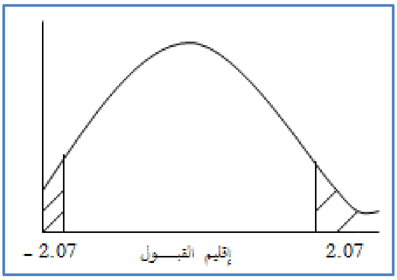
ثانياً:- الاختبار : لحساب قيمة الاختبار نحتاج للانحراف المعياري المرجح Sp :

$$Sp = \sqrt{\frac{(12-1)10.2+(12-1)6.2}{12+12-2}} = 2.86$$

$$T = \frac{31.75-29.5}{2.86\sqrt{\frac{1}{12}+\frac{1}{12}}} = 1.93$$

ثالثاً:- الحدود الحرجة والمقارنة والاستنتاج :

قيمة T المعيارية المقدرة عند α 0.05 وdf =22 تساوي 2.075 و – 2.075



وحيث أن T الاختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأنه لا توجد أفضلية قاطعة أو مؤكدة بين الطريقتين .

2- أستخدم سماد النيتروجين في 10 قطع من الأرض وكان متوسط إنتاج القطعة 82.5 قنطارا وانحرافها المعياري 10 قناطير وأستخدم سماد الفوسفات في 15 قطعة أخرى فكان متوسط إنتاجها 90.5 قنطارا وانحرافها المعياري 20 قنطارا . استخدم α 0.05 لاختبار هل يوجد اختلاف بين مفعول السمادين يجعلنا نفاضل بينهما .

الحل :

أولاً:- المعطيات والفرضيات :

الأول ... الثاني

$m = 10 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots n = 15$

$\overline{X} = 82.5 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \overline{Y} = 90.5$

$S_1^2 = 100 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots S_2^2 = 400$

$H0 : \mu 1 = \mu 2$

$HA : \mu 1 \# \mu 2$

$\alpha = 0.05$

$df = 10 + 15 - 2 = 23$

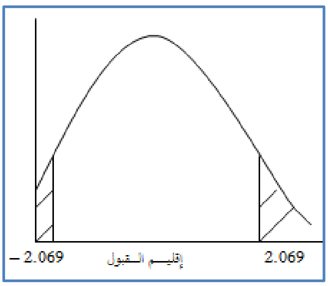
$$Sp = \sqrt{\frac{(10-1)100 + (15-1)400}{23}} = 16.8$$

ثانياً:- الاختبار :

$$T = \frac{82.5 - 90.5}{16.8\sqrt{\frac{1}{10} + \frac{1}{15}}} = -1.16$$

ثالثاً:- الحدود الحرجة والمقارنة والاستنتاج :

القيمة المعيارية المقدرة عند مستوى ثقة 95% و23 درجة حرية تساوي 2.069 و -2.069



وبما أن T ليست بالإقليم الحرج نقبل H0 ونستنتج أنه ليس هناك ما يدل على أن مفعول السمادين مختلف اختلافا كبيرا يجعلنا نفاضل بينهما.

تمرينات:

1- من المعتقدات الراسخة لدى الجغرافيين الإنجليز هو أن أسر عمال مناجم الفحم أكبر من أسر عمال القطن بإقليم فكتوريا البريطانية، البيانات التالية هي لعينتين من أسر عمال الفحم وعمال القطن، هل تؤيد رأي الجغرافيين، $\alpha = 0.05$

عمال الفحم .... عمال القطن

$m = 213$..............................$n = 250$

$\overline{X} = 3.46$.............................$\overline{Y} = 3.14$

$S_1^2 = 0.02$............................$S_2^2 = 0.016$

2- باحث جغرافي يرغب في اختبار الفكرة السائدة بأن أسعار المنازل بالمدن الكبيرة تختلف عن أسعار المنازل بالمدن الصغيرة وتحصل على البيانات التالية، عند معدل ثقة 95% هل هناك اختلاف جوهري.

المدن الكبيرة .... المدن الصغيرة

$m = 12$............................$n = 10$

$\overline{X} = 27164$.....................$\overline{Y} = 22732$

$S_1^2 = 17627$....................$S_2^2 = 12588$

3- أجريت دراسة لاختبار الفرق بين رواتب أساتذة الكليات العامة وأساتذة الكليات الخاصة بأحد الولايات الأمريكية وكانت البيانات كما يلي:

الكليات العامة .... الكليات الخاصة

$m = 100$...........................$n = 200$

$\overline{X} = 15000$......................$\overline{Y} = 14000$

$S_1^2 = 1300$.......................$S_2^2 = 1400$

أختبر عند مستوى ثقة 95% هل رواتب أساتذة الكليات العامة فعلا أكبر.

4- خصصت الدولة منحا لأقسام المحاصيل في 10 كليات للزراعة لفحص نوعين جديدين من البذور المحسنة، وتم زراعة النوعين في مساحتين متساويتين في كل كلية، وجاءت النتائج (كج/مساحة) كم يلي:

الكلية	1	2	3	4	5	6	7	8	9	10
النوعA	28	23	35	41	44	29	37	31	38	40
النوعB	45	25	31	38	50	33	36	40	43	44

أوجد هل هناك اختلاف جوهري بين إنتاجية النوعين، α=0.05

5- فيما يلي بيانات عن سجل كبريات المطر اليومي المسجلة بمحطة أرصاد مدينة نالوت بالجبل الغربي وهي مقسمة إلى فترتين الأولى من عام 1945- 1970 والثانية للفترة 1971- 1996 اختبر هل هناك تغير دال إحصائيا في المطر يؤيد فكرة التغير المناخي وفي أي اتجاه كان التغير. (مقيلي، 1924، ص97)

معطيات الفترة الأولى:

مجموع المطر خلال الفترة الأولى = 808.9 ملم

متوسط المطر للفترة الأولى $\bar{X}_1$ = 31.11 ملم

المجموع التربيعي للفرق بين قيم المطر والمتوسط $\sum(X-\bar{X}_1)^2$= 8859.49ملم

الانحراف المعياري للفترة الأولى= 18.82

معطيات الفترة الثانية:

مجموع المطر خلال الفترة الثانية = 1181.5 ملم

متوسط المطر للفترة الثانية $\bar{X}_2$ = 45.44 ملم

المجموع التربيعي للفرق بين قيم المطر والمتوسط $\sum(X-\bar{X}_2)^2$ = 11645.31 ملم

الانحراف المعياري للفترة الثانية = 21.58

3- اختبار الاختلاف بين نسبتين:

كثيرا ما يهمنا معرفة هل الاختلاف بين نسبتي عينتين ناتجة عن أخطاء في التعيين أم عن حقيقة أن هاتين النسبتين مختلفتين أساسا في المجتمعات الأصلية التي أخذت منهما العينتان. **مثلا،** أعطي العقار A لعينة عشوائية من 200 رجل وأعطي العقار B لعينة أخرى من 300 رجل وكانت نتائجهم كما يلي:

الحالـــة	العقارA	العقارB
عدد المصابين بعد تناولهم العقار	80	105
عدد غير المصابين بعد تناول العقار	120	195
المجمـــــــوع	**200**	**300**

لاحظ ما يلي:

1- من مجموع 200 رجلا أخذوا العقار A أصيب 80 منهم، أي بنسبة 80\200 =0.40

2- من مجموع 300 رجلا أخذوا العقار B أصيب 105منهم، أي بنسبة 0.35=300\105

3- الفرق بين النسبتين 0.40-0.35 =0.05 ، والسؤال الذي نود الإجابة عليه هو هل هذا الاختلاف بين النسبتين كبير يدل على وجود اختلاف جوهري بين مفعول العقارين أو أنه اختلاف قليل جدا وناتج عن أخطاء التعيين ولا يجب الاهتمام به . عموما، إذا كانت P1 و P2 يمثلان نسبتن لعينتين مختارتين من مجتمعين، فنحن نرغب في اختبار الفرضية التي تنص على أن النسبة الأولى P1 لا تختلف عن النسبة الثانية P2 بل هي مساوية لها ضد الفرضية البديلة القائلة بوجود اختلاف:

H0:P1=P2

HA:P1#P2

والاختبار الإحصائي هو:

$$Z = \frac{P1 - P2}{\sqrt{\hat{P}\left(1-\hat{P}\right)\left(\frac{1}{m}+\frac{1}{n}\right)}} \quad \text{...(9)}$$

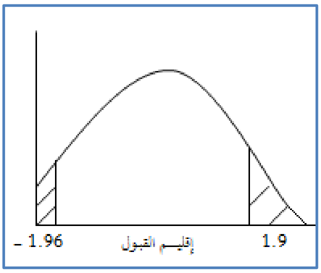
$\hat{P}$ = النسبة المرجحة من العينتين وتحسب بالمعادلة التالية :

$$\hat{P} = \frac{P1 + P2}{2}$$

ومن المثال نجد أن P1 = 0.40 وP2 = 0.35 ومن ثم فإن النسبة المرجحة = 0.375 الآن نحسب Z الاختبارية كما يلي :

$$Z = \frac{0.40 - 0.35}{\sqrt{0.375(1-0.375)\left(\frac{1}{200}+\frac{1}{300}\right)}} = 1.134$$

أما قيم $Z_{\frac{\alpha}{2}}$ = 1.96 و- 1.96



وبما أن Z واقعة ضمن حدود الثقة نقبل H0 ونستنتج بأنه ليس هناك اختلاف جوهري بين مفعول العقارين.

مثال 2- لقد تم رش صندوق يحتوي على 120 ذبابة بمبيد حشري من النوع A مات منها 95 ذبابة، ورش صندوق آخر يحتوي على 145 ذبابة بمبيد من النوع B مات منها 124 ذبابة. استخدم $\alpha = 0.05$ لاختبار هل هناك اختلاف كبير بين مفعول المبيدين.

الحل:

أولاً:- المعطيات والفرضيات:

المبيد A المبيد B

$$m = 120 \ldots\ldots\ldots\ldots\ldots\ldots\ldots n = 145$$

$$P1 = \frac{95}{120} = 0.7916 \ldots\ldots\ldots\ldots P2 = \frac{124}{145} = 0.8551$$

$$\hat{P} = \frac{0.7916 + 0.8551}{2} = 0.826$$

$$H0 : P1 = P2$$

$$HA : P1 \# P2$$

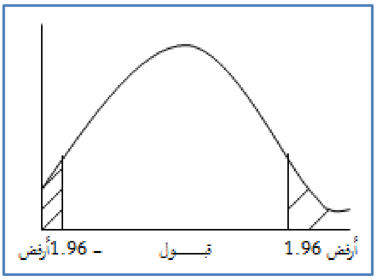
$$\alpha = 0.05$$

ثانياً:- الاختبار:

$$Z = \frac{0.7916 - 0.8551}{\sqrt{0.826(1 - 0.826)\left(\frac{1}{120} + \frac{1}{145}\right)}} = -1.35$$

ثالثاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج:

قيمة Z المعيارية لاختبار الطرفين = 1.96 و − 1.96



Z ليست بالإقليم الحرج لذلك نقبل H0 ونستنتج بأن مفعول المبيدين ليس مختلفا اختلافا ذو دلالة إحصائية.

تمرينات:

1- لقد تم سؤال عينتين من أساتذة وطلبة الجامعات وطلبة التعليم المتوسط الأولي من الذين أعمارهم أقل من 40 سنة وعددهم 56 فردا صنف 74% منهم بأن المشكلة تعتبر خطرا داهما بينما العينة الثانية من الذين أعمارهم فوق 41 سنة متكونة من 35 فردا كانت نسبة من صنفوا المشكلة بالخطر الداهم 52%. أختبر الفرضية القائلة بأن صغار السن أكثر تأثرا بالمشكلة من كبار السن.

2- تشير بعض الدراسات إلى أن مضغ التبغ يزيد من سرعة تجلط الدم والإجهاض لدى الحوامل، فالنساء اللاتي يمضغون التبغ يجهضن 30 لكل 1000 حالة ولادة مقابل 17 لكل 1000 حالة لغير الماضغات، هل نسبة الإجهاض لدى الماضغات أكبر كثيرا من نسبتها لدى غير الماضغات، α=0.05 .

3- من دراسة عينة متكونة من 500 رجلا مدخنا تبين أن 50% منهم تعرضوا للطلاق ومن دراسة عينة أخرى متكونة من 400 رجلا من غير المدخنين تبين أن 32% منهم تعرضوا للطلاق عند مستوي ثقة 95% هل للتدخين دور في زيادة حالات الطلاق.

4- لمعرفة أثر المناخ على غياب الطلبة عن الدراسة، أجريت دراسة على عينتين، الأولى من ولاية مين (باردة شتاء) والثانية من ولاية ألباما (دافئة شتاء) وكانت النتائج كالتالي:

الحالة	مين	الباما
غياب على الأقل يـوم واحد	72	70
المجمـوع	**300**	**400**

استخدم α 0.05 لاختبار هل المناخ البارد يؤدي إلى غياب أكثر بين الطلبة.

5- أجري استبيان لمقارنة تكنولوجيا المنازل بين المدينة وضواحيها فوجد أن 63 امرأة من كل 100 بالمدينة يمتلكن مجففا للشعر وبالمقابل 59 امرأة من بين كل 125 بالضاحية يمتلكن مجففا، فهل يوجد ما يدل على أن هناك اختلاف جوهري بين المدينة والقرية في نسبة امتلاك التكنولوجيات، استخدم نسبة ثقة 95%.

6- سئل أفراد عينتين عشوائيتين الأولى متكونة من 100 رجل والأخرى من 100 امرأة عن امتلاكهم سيارة أم لا أجاب 31 رجل و 24 امرأة بنعم، فهل يمكن استنتاج بأن الرجال أكثر امتلاكا للسيارات، α 0.05 .

7- أظهرت دراسة لعينة من 900 رجل وأخرى من 700 امرأة أجنبية في فرنسا أن 80% من الرجال و 75 % من النساء يرغبون في المشاركة السياسية بالتصويت في الانتخابات المحلية، أوجد هل هناك فرق كبير بين رغبة الرجال والنساء الأجانب في المشاركة السياسية، α 0.01

8- باحث يريد التحقق من دور الطقس في غياب الطلاب عن المدارس، فأخذ عينة من 440 طالبا ووجد أن عدد الغائبين منهم في أحد أيام الشتاء 30 طالبا مقارنة بعدد 20 طالبا في أحد أيام الصيف، هل هناك ما يؤيد أن الغياب في الشتاء أكثر، α 0.05

9- يوجد افتراض بأن الطلاب الذين يجمعون بين العمل والدراسة أقل استيعابا للدروس ومن ثم فإن نتائجهم أضعف في الامتحانات، أخذت عينة من 80 طالبا من الذين يعملون و70 من المتفرغين للدراسة فتبين أن عدد الناجحين 60 من الذين يعملون و 63 من المتفرغين، استخدم α 0.05 لاختبار صحة الافتراض من عدمه.

10- بلغ عدد مواليد سنة 1990 بأحد المدن 760 طفلا منهم 384 ذكرا و 376 أنثى وبالمقابل بلغ عدد مواليد سنة 1995بنفس المدينة 760 طفلا منهم 384 ذكرا و 376 أنثى، عند مستوى ثقة 95% أوجد:

أ- هل هناك اختلاف جوهري بين نسبتي ذكور الفترتين؟.

ب- هل هناك اختلاف جوهري بين نسبتي إناث الفترتين؟.

11- أجري استبيان لعدد 200 طالبا جامعيا فوجد أن 150 منهم يفضلون اتخاذ إجراءات صارمة للحفاظ على البيئة، وبالمقال وجد أن 205 من أصل 300 خريج جامعي يفضلون الإجراءات الصارمة، عند معدل ثقة 0.95 هل طلبة الجامعة أكثر ميلا للإجراءات الصارمة.

12- فصل دراسي متكون من 200 طالبا قسم إلى مجموعتين 100 لكل منهما، أعطيت المجموعة الأولى مصل مضاد للرشح بينما تركت الثانية بدونه. لقد وجد فيما بعد أن عدد الذين أصيبوا بالرشح بلغ 45 في المجموعة الأولى و 55 في المجموعة الثانية. هل هناك ما يدل على أن المضاد فعال في تقليل الإصابة، α =0.05

الخلاصة:

من اختبارات المجتمع الواحد والمجتمعين الإحصائيين يمكن الخروج بالخلاصة التالية:

أولاً:- أدرس المشكلة جيداً وقرر هل هي تخص المجتمع الواحد أو المجتمعين.

أ- اختبار المجتمع الواحد:

إذا كانت المشكلة تخص المجتمع الواحد في زمنين مختلفين أو في زمن واحد لكن باختلاف المؤثر الخارجي، فأمامنا تطبيق أحد الاختبارات الأربعة التالية بناءً على نوع المعلم المراد اختباره:

1- إذا كان المعلم قيد الاختبار هو المتوسط الحسابي $\overline{X}$ فأمامنا اختبار Zأو T.

*- إذا كان تباين المجتمع σ^2 معلوما نستخدم اختبار Z:

$$Z = \frac{\overline{X} - \mu 0}{\frac{\sigma}{\sqrt{n}}}$$

*- إذا كان تباين المجتمع غير معلوم نستخدم اختبارT :

$$T = \frac{\overline{X} - \mu 0}{\frac{S}{\sqrt{n}}}$$

2- أما إذا كان المعلم المراد اختباره هو التباين σ^2 نستخدم كاي تربيع:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0}$$

3- وإذا كان المعلم هو النسبة المئوية P نستخدم اختبار Z:

$$Z = \frac{P1 - P2}{\sqrt{\hat{P}\left(1-\hat{P}\right)\left(\frac{1}{m}+\frac{1}{n}\right)}}$$

ب- اختبار مجتمعين إحصائيين:

في حلة ما أتضح أن المشكلة تخص مجتمعين إحصائيين في آن واحد، فأمامنا ثلاثة اختيارات:

1- إذا كان معلم الدراسة هو المتوسط الحسابي للعينة الأولى $\overline{X}$ والمتوسط الحسابي للعينة الثانية $\overline{Y}$ فأمامنا اختبارين:

*- إذا كان تبايني المجتمعين متوفرين نستخدم Z:

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$$

*- إذا لم يوجد تبايني المجتمعين نستبدلهما بتبايني العينتين ونستخدم اختبار T:

$$T = \frac{\overline{X} - \overline{Y}}{Sp\sqrt{\frac{1}{m}+\frac{1}{n}}}$$

$$Sp = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}}$$

2- إذا كان معلم الدراسة هو النسبة نستخدم اختبار Z:

$$Z = \frac{P1 - P2}{\sqrt{\hat{P}\left(1-\hat{P}\right)\left(\frac{1}{m}+\frac{1}{n}\right)}}$$

$$\hat{P} = \frac{P1 + P2}{2}$$

الفصل الرابع
تحليل الارتباط والانحدار الخطي

تمهيد:

يهتم تحليل الارتباط Correlation Analysis بدراسة نوع العلاقة بين متغيرين Y,X أو أكثر عن طريق الوصول إلى معامل رقمي يوصف هذه العلاقة، وعن طريق هذا المعامل الرقمي يتسنى للباحث القيام بإجراء أحكام أو تنبؤات عن أحد المتغيرات بفضل ما يعرفه عن المتغير الآخر. إذا موضوع الارتباط مهتم أساسا بدراسة العلاقة بين متغيرين أو أكثر من حيث نوع العلاقة أو قيمتها الحسابية، أي بنوع التغير المشترك بين متغيرين يرمز لهما Y,X.وهذه العلاقة قد تكون سببية بحيث يمكننا أن نميز بوضوح بين السبب والنتيجة. ففي المناطق شبه الجافة التي تعتمد فيها الزراعة والرعي على المطر نجد أن الوضع الاقتصادي بصفة عامة يزدهر خلال السنوات المطيرة ويتدهور خلال فترات الجفاف، وفي مناطق التربة المجهدة توجد علاقة بين كميات الأسمدة المضافة إلى التربة وبين كميات الإنتاج الزراعي، وهناك علاقة بين عدد ساعات المذاكرة ودرجات الطالب في الامتحان. إن العلاقة في هذه الحالات لا تخرج عن سبب ونتيجة، فزيادة المطر والسماد يسبب زيادة الإنتاج الزراعي، وزيادة ساعات المذاكرة يؤدي إلى تحسن درجات الامتحان...الخ.

وفي دراستنا لتحليل الانحدار قد نحصر العلاقة بين متغيرين فقط ويسمى التحليل في هذه الحالة بالانحدار البسيط، وقد نحصرها بين عدد من المتغيرات في آن واحد كالعلاقة بين زيادة الإنتاج الزراعي والمطر والسماد والمبيدات ورأس المال والسوق فالإنتاج مرتبط ليس بعامل واحد وإنما بالعديد منها. وهذا النوع من التحليل هو الشائع في الواقع، لكن لتسهيل الموضوع للفهم سوف نبدأ بدراسة الانحدار البسيط ثم ننتقل للانحدار المتعدد لاحقا.

في العلوم التطبيقية والاجتماعية كثيرا ما يهتم الباحث بدراسة العلاقة بين متغيرين اثنين Y,X كعمر الطفل ودرجة ذكائه، وبين كمية المطر وكمية المحصول، وبين عدد السيارات ومعدل الحوادث على الطرقات، وبين كثافة الغطاء النباتي ومعدل الانجراف وبين الكثافة السكانية ومعدل

الجريمة. فالعلاقة في هذه الأحوال تسمى بالارتباط البسيط لأنها لا تتعدى البحث عن العلاقة بين متغيرين هما Y,X.

افرض أن لدينا عينة متكونة من عدد n من المشاهدات لكل من Y,X، فإن العلاقة بينهما لا تخرج في شكلها عن النماذج التالية:

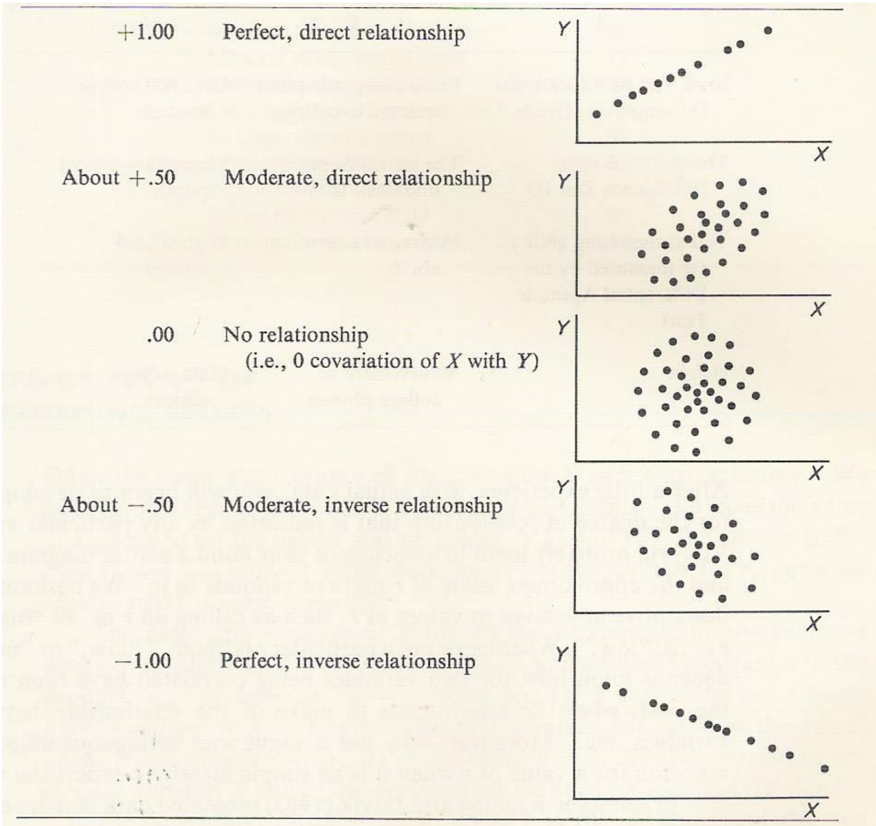
1- الارتباط الخطي الكامل:- إذا كانت البيانات موزعة بحيث يبدأ الخط من أسفل اليسار وينتهي بأعلى اليمين وجميع نقاط توزيع Y,X واقعة على خط الانحدار فالعلاقة توصف بأنها طردية كاملة الارتباط، وأي تغير في قيمةX سوف يؤدي إلى تغير مماثل (بنفس النسبة) في قيمة Y ، والتغير في قيمة Y يتبع نفس اتجاه التغير في قيمة X ، كالعلاقة بين سرعة الرياح وبين ارتفاع أمواج البحر، وبين معدل التبخر ودرجة الحرارة، وبين تقدم العمر وزيادة ضغط الدم، فهنا العلاقة دائما موجبة.

أما إذا كانت البيانات موزعة بحيث ترسم خطا هابطا من أعلى اليسار إلى أسفل اليمين وجميع النقاط واقعة على الخط فالعلاقة توصف بأنها عكسية كاملة الارتباط، والتغير في قيمة X يؤدي إلى تغير في قيمة Y بنفس النسبة لكن في الاتجاه المعاكس، كالعلاقة بين الارتفاع عن مستوى سطح البحر وتناقص درجة الحرارة، وبين الارتفاع وتناقص الضغط الجوي وبين طول المسافة التي تقطعها السيارة وكمية الوقود المتبقي في الخزان.

2- الارتباط الخطي غير الكامل: إذا كانت نقاط توزيع Y,X في الشكل البياني لا تتبع خط الانحدار المستقيم وإنما تبدو مبعثرة حوله فإن العلاقة في هذه الحالة ليست كاملة وتتباين قيمتها من قوية إذا كانت النقاط قريبة من الخط، وضعيفة إذا كانت شديدة التبعثر حوله، وتكون العلاقة صفرية إذا كانت النقاط مبعثرة بطريقة لا يعرف لها اتجاه صاعد ولا هابط. وعلى ذلك ينصح دائما بتوقيع أزواج القيم(Y,X) على ورقة رسم بياني من أجل الحصول على شكل الانتشار الذي يعطي فكرة مبدئية وسريعة عن نوع العلاقة بين المتغيرين. ويجب التأكد دائما من التحديد الصحيح للمتغير الحر X والمتغير المعتمد أو التابع Y، وذلك بتوزيع مدايات المتغير الحر X على المحور الأفقي ومدايات المتغير التابع على المحور الرأسي.

ويطلق على المعامل الذي يوصف نوع العلاقة بين متغيرين معامل الارتباط ويرمز له بالحرف r وتنحصر قيمته بين+1 أو -1 والارتباط الكامل نادر الوجود، ولكن نعتبر الارتباط قوي إذا كانت قيمة r كسر عشري قريب من +1 أو -1 . فعلى سبيل المثال، نعتبر r (0.80 أو

0.90 أو 0.95) ارتباط قوي، ويضعف الارتباط مع اقتراب قيمة r من الصفر سواء من الاتجاه السالب أو الموجب.



Gene V.Glass & Jullian C Stanley " Statistical Methods in Education And Psychology" Prentice-Hall, Inc., New Jersey 1970.p117

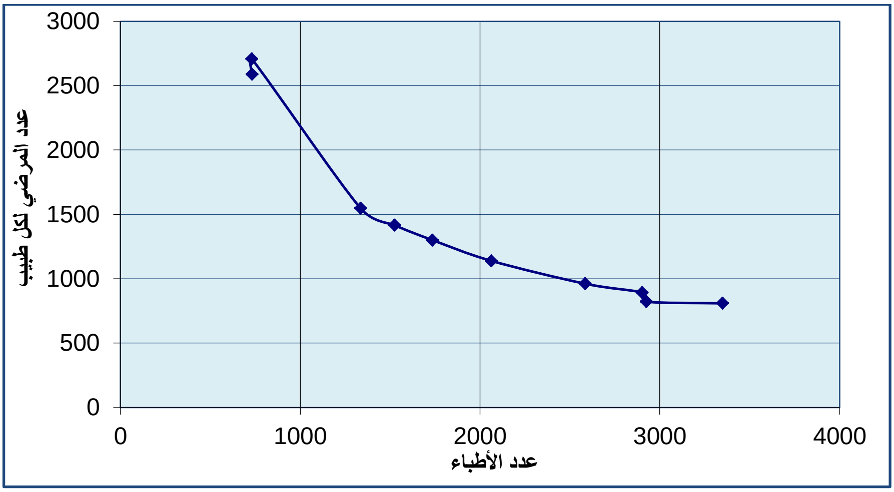
مثال:

1- الجدول التالي يبين التطور الحاصل في أعداد الأطباء الليبيين من 733 طبيبا عام 1969 إلي 3350 عام 1978 وارتبط بذلك بالضرورة تناقص في أعداد المرضى الذين يخدمهم كل طبيب من 2588 مريض عام 1969 إلى 809 مريض عام 1978. أرسم تخطيط الانتشار للحصول على صورة أوضح لهذه العلاقة.

الجدول (19) التطور الحاصل بين عدد الأطباء وأعداد المرض الليبيين للفترة 1969 – 1978.

السنة	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
عدد الأطباء	733	731	1336	1525	1736	2063	2585	2903	2926	3350
مريض لكل طبيب	2588	2707	1547	1416	1300	1139	961	892	822	809

شكل (11) تطور العلاقة بين الأطباء والمرضى الليبيين خلال الفترة 1969–1978.



تمرينات:

1- الجدول التالي يبين العلاقة بين الارتفاع عن مستوى سطح البحر وكل من درجة حرارة الهواء والضغط الجوي. فالبيانات توضح أن هناك تناقصا في الضغط والحرارة مع زيادة الارتفاع في طبقة التروبوسفير. ارسم تخطيط الانتشار مرة للارتفاع والحرارة وأخرى للارتفاع والضغط لكي تتضح الصورة أكثر، ثم أذكر السبب في تناقص الحرارة والضغط بالارتفاع.

جدول(20) العلاقة بين الإرتفاع وكلا من درجة الحرارة والضغط الجوي.

الضغط MB	درجة الحرارة C	الارتفاع M
1013.25	15.2	0
898.76	8.7	1000
795.01	2.2	2000
701.21	-4.3	3000
616.60	-10.8	4000
540.48	-17.3	5000
265.00	-49.7	10000

H. Critshfield. General Climatology,P86.

2- الغازات تذوب في المياه الباردة وتتشبث بالجليد، لكن بارتفاع الحرارة تتطاير من المياه ويقل تركيزها به، الجدول التالي يبين تناقص تركيز الأكسجين مع ارتفاع درجة حرارة الماء.

جدول (21) العلاقة بين درجة الحرارة وتركيز الأكسجين بالماء.

درجة الحرارة	0	5	10	15	20	25	30
تركيز الأكسجين ppm	14.6	12.7	11.3	10.1	9.1	8.3	7.5

المصدر: الهادي أبو لقمة و محمد الأعور، **الجغرافية البحرية**. الدار الجماهيرية للنشر والإعلان والتوزيع، مصراتة، 1999، ص147.

3- الجدول التالي يبين تطور الحركة السياحية الدولية للفترة 1950-1990. أرسم تخطيط الانتشار.

الجدول(22) يبين تطور الحركة السياحية الدولية للفترة 1950-1990

السنة	1960	1965	1970	1975	1980	1985	1990
عدد السياح بالملايين	79.296	112.729	159.690	214.357	284.841	321.240	443.477

المصدر: فضل أحمد يونس، **الجغرافيا السياحية**، دار النهضة العربية، بيروت، 1993، ص17.

حساب معامل الارتباط:

تحسب قيمة معامل ارتباط بيرسون بالمعادلة التالية:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\left[\sum X^2 - \frac{\left(\sum X\right)^2}{n}\right]}\sqrt{\left[\sum y^2 - \frac{\left(\sum y\right)^2}{n}\right]}} \quad ....(10)$$

متال1- تحصل باحث على بيانات عن حجم الأسرةX وحجم نفقاتهاY بالدنانير، والمطلوب حساب قيمة معامل الارتباط بين حجم الأسرة وحجم نفقاتها.

جدول(23) العلاقة بين حجم الأسرةX وحجم نفقاتهاY بالدنانير

حجم الأسرة X	1	3	4	6	8	9	11	14
النفقات آلاف الدنانيرY	1	2	4	4	5	7	8	9

الحل:

أولاً:- نكون جدول لحساب عناصر المعادلة كما يلي:

X	X^2	Y	Y^2	XY
1	1	1	1	1
3	9	2	4	6
4	16	4	16	16
6	36	4	16	24
8	64	5	25	40
9	81	7	49	63
11	121	8	64	88
14	196	9	81	126

$$\sum X = 56 \ldots\ldots\ldots\ldots \sum X^2 = 524 \ldots\ldots\ldots\ldots\ldots\ldots \sum Y = 40 \ldots\ldots\ldots\ldots\ldots \sum Y^2 = 256$$

$$\sum XY = 364$$

ثانياً:- نعوض في المعادلة رقم (1) ونحسب قيمة معامل الارتباط r :

$$r = \frac{364 - \frac{(56)(40)}{8}}{\sqrt{\left[524 - \frac{(56)^2}{8}\right]}\sqrt{\left[256 - \frac{(40)^2}{8}\right]}} = 0.97$$

ثالثاً:- الاستنتاج: معامل الارتباط r = 0.97 يعني أن هناك علاقة موجة وقوية.

مثال 2- فيما يلي أسعار الجملة لبعض المواد الغذائية بالدولار خلال يومين من عامي 1971و1972. أوجد قيمة الارتباط بين أسعار السلع.

جدول(23) أسعار الجملة لبعض المواد الغذائية

بالدولار خلال يومين من عامي 1971و1972.

نوع السلعة	X	Y
دقيق	6.65	6.96
قهوة	0.44	0.45
كوكا	0.28	0.27
سكر	0.130	0.125
زبدة	0.69	0.71
بيض	0.34	0.32
لحم ضأن	23.80	17.25
لحم بقر	35.25	31.50
دجاج	4.50	4.30

الحل:

أولاً:- نكون جدول لحساب عناصر معادلة الارتباط كما يلي:

X	X^2	Y	Y^2	XY
6.96	48.4416	6.65	44.2225	46.284
0.45	0.2025	0.44	0.1936	0.198
0.27	0.0729	0.28	0.0784	0.0756
0.125	0.0156	0.130	0.0169	0.01625
0.71	0.5041	0.69	0.4761	0.4899
0.32	0.1024	0.34	0.1156	0.1088
17.25	297.5625	23.8	566.44	410.55
31.50	992.25	35.25	1242.5625	1110.375
4.30	18.49	4.90	34.01	21.07

$$\sum X = 61.885............\sum X^2 = 1357.6416....................\sum Y = 72.48$$
$$\sum Y^2 = 1878.1156......................\sum XY = 1589.1676$$

ثانياً:- نعوض في المعادلة رقم (1) لنحسب معادلة الارتباط:

$$r = \frac{1589.1676 - \frac{(61.9)(72.5)}{9}}{\sqrt{\left[1357.6 - \frac{(61.9)^2}{9}\right]}\sqrt{\left[1878.1 - \frac{(72.5^2)}{9}\right]}} = 0.99$$

ثالثاً:- الاستنتاج الرقم 0.99 دليل على أن الارتباط موجب وقوي.

اختبار الدلالة الإحصائية لمعامل الارتباط:

لاختبار الدلالة الإحصائية لمعامل ارتباط بيرسون r نستخدم الاختبار الثائي كما يلي:

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

حيث تقارن قيمة t الاختبارية بقيمة t_α بدرجات الحرية df = (n-2). وإذا كان حجم العينة n أكبر من 25 يمكن استخدام اختبار Z كما يلي:

$$Z = r\sqrt{n-2}$$

حيث تقارن قيمة Z الاختبارية بقيمة Z_α ، أما الفرضيات الاختبارية فهي كما يلي:

H0 : r = 0

HA :r # 0

مثال:

أستخدم اختبار t للتأكد من الدلالة الإحصائية لمعامل الارتباط المحسوب لأسعار السلع الغذائية ($\alpha = 0.05$).

الحل:

أولاً:- الفرضيات:

H0 : r = 0

HA :r # 0

ثانياً:- نحسب قيمة t الاختبارية:

$$t = 0.99\sqrt{\frac{9-2}{1-0.99^2}} = 18.56$$

ثالثاً:- قيمة $t_{(0.05,7)}$ تساوي 2.365

رابعاً:- المقارنة والاستنتاج: بما أن t أكبر من 2.365(مما يعني أنها واقعة بالإقليم الحرج) نرفض H0 ونستنتج بأن هناك ارتباط دو دلالة إحصائية بين أسعار سلع الفترتين.

تمرينات:

1- الجدول التالي يبين عدد الوفيات من المدنيين وتقديرات الخسائر المادية من الحرائق بالولايات المتحدة خلال الفترة 1978-1994 والمطلوب حساب قيمة ارتباط بيرسون بين الوفيات والأضرار المادية والتعليق على النتائج.

جدول(24) الوفيات وخسائر الممتلكات جراء الحرائق المسجلة بالولايات المتحدة خلال الفترة 1978- 1994.

السنة	الوفيات	الأضرار بلايين الدولارات
1978	8130	4008
1979	7780	4851
1980	6505	5579
1981	6700	5625
1982	6020	5894
1983	5920	6320
1984	5240	7602
1985	6185	7753
1986	5850	8488
1987	5810	8504
1988	6215	9626
1989	5410	9514
1990	5195	9495
1991	4465	11302
1992	4730	13588
1993	4635	11331
1994	4275	12333

المصدر: باتريك . ل . أبوت، ترجمة: توفيق علي منصور، الكوارث الطبيعية، المجلد الثاني، المجلس الأعلى للثقافة، القاهرة، 2003م، ص 372.

2- الجدول التالي يبين عدد الوفيات من المدنيين وتقديرات الخسائر المادية من الفيضانات بالولايات المتحدة خلال الفترة 1969-1989 والمطلوب حساب قيمة ارتباط بيرسون بين الوفيات والأضرار المادية والتعليق على النتائج.

جدول(25) الوفيات وخسائر الممتلكات جراء الفيضانات المسجلة بالولايات المتحدة خلال الفترة 1969- 1989.

السنة	الوفيات	خسائر الممتلكات
1969	297	903
1970	135	225
1971	74	288
1972	540	3449
1973	105	856
1974	121	576
1975	114	1051
1976	187	1000
1977	212	1393
1978	120	1000
1979	100	4000
1980	97	1500
1981	90	1000
1982	155	3500
1983	200	4100
1984	126	4000
1985	304	3000
1986	80	4000
1987	82	1490
1988	29	114
1989	81	415

المصدر: باتريك . ل . أبوت، ترجمة: توفيق علي منصور، الكوارث الطبيعية، المجلد الثاني، المجلس الأعلى للثقافة، القاهرة، 2003م، ص 372.

3- الجدول التالي يبين عينة لأعمار 12 زوج، والمطلوب حساب معامل ارتباط بيرسون لمعرفة هل هناك علاقة بين عمر الزوجة وعمر الزوج.

جدول(26) عينة لأعمار 12 زوج.

عمر الزوج	37	45	26	49	36	35	46	29	25	81	44	22
عمر الزوجة	32	29	20	52	20	30	39	22	22	52	35	20

تحليل الانحدار الخطي:

إن الهدف من تحليل الارتباط الخطي correlation Analysis هو إيجاد علاقة خطية بين المتغيرات، وإذا وجدت هذه العلاقة فإن معامل الارتباط (r) سوف يخبرنا عن قرب هذه العلاقة من الخط المستقيم. أما نظرية الانحدار Regression فهي توصف هذه العلاقة الخطية بلغة كمية. وبمجرد أن تحسب المعادلة التي توصف العلاقة يمكن استعمالها في إجراء تنبؤات بخصوص أحد المتغيرات من خلال معرفتنا لخصائص المتغير الآخر.

والمقصود بالعلاقة الخطية هي أن العلاقة بين المتغيرين Y,X يمكن أن توصف بواسطة خط مستقيم، ولتوضيح ذلك لا بد من سياق المثالين التاليين:

مثال 1– افرض أن سائق سيارة أجره يشترط 0.5 دينار للكيلومتر الواحد، وافرض كذلك أن X تمثل عدد الكيلومترات المتنقلة و Y تمثل مجموع النقود التي سيتقاضاها السائق، فبالإمكان صياغة تلك العلاقة في صورة المعادلة التالية:

$$Y=0.5X$$

والآن، افرض أن سائق سيارة الأجرة أخد يشترط 2 دينار لمجرد أن يدور مفتاح المحرك تم 0.5 دينار لكل كيلومتر، ففي هذه الحالة تتغير صورة المعادلة السابقة إلى الشكل التالي:

$$Y=2+0.5X$$

إذا، لكل عدد من الكيلومترات المتنقلة يمكن حساب مجموع النقود Y التي سيتقاضاها السائق كما في الجدول التالي:

الأجرة (دنانير)	عدد الكيلومترات(X)
2.0	0
2.5	1
3.0	2
3.5	3
4.0	4
4.5	5
5.0	6
5.5	7
6.0	8

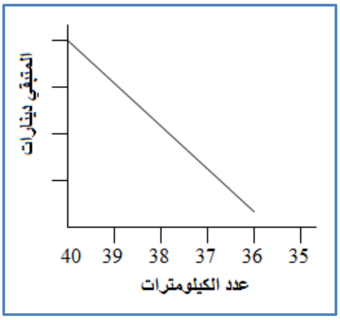
وبعد حساب قيم Y,X بهذا الشكل يمكن توقيعها على ورقة رسم بياني، وسيتضح من الشكل أن جميع النقاط تقع على خط مستقيم صاعد تمثله المعادلة السابقة.

مثال 2– افرض أن شخصا عنده 40 دينارا في جيبه، استأجر سيارة بمبلغ 0.5 دينار للكيلومتر، وإذا فرضنا أن X تمثل عدد الكيلومترات المتنقلة و Y تمثل النقود المتبقية معه بعد تنقل عدد X من الكيلومترات. في هذه الحالة يمكن صياغة هذه العلاقة في المعادلة التالية:

$$Y=40-0.5X$$

إن قيم Y لمختلف قيم X مبوبة في الجدول التالي، ومن الشكل التالي يتبين أن جميع نقاط Y,X تقع على خط مستقيم نازل من أعلى اليسار إلى أسفل اليمين.

الدينارات المتبقية	عدد الكيلومترات
40	0
39	2
38	4
37	6
36	8
35	10



وفي العموم يمكن صياغة علاقة الخط المستقيم بين متغيرين Y,X بالمعادلة التالية:

$$Y=a+bX \quad (11)$$

حيث أن :

a= قيمة Y عندما تكون X=0، كما يمكن تعريفها بنقطة تقاطع خط الانحدار مع Y.

b= درجة ميل خط الانحدار، إنها تحدد قيمة التغير الحاصل في Y الناتج عن تغير X وحدة واحدة. فإذا كانت b موجبة الإشارة، فإن الخط المستقيم يصعد من أسفل اليسار إلى أعلى اليمين، أما إذا كانت سالبة الإشارة فإن الخط ينزل من أعلى اليسار إلى أسفل اليمين.

كيفية رسم خط الانحدار:

من معرفة قيمة r يصير بالإمكان الحكم على وجود علاقة خطية بين Y,X أم لا، والحكم كذلك على درجة هذه العلاقة قوية أو ضعيفة، لكن معامل الارتباط لوحده لا يكفي ولا يمكن استعماله في إجراء تنبؤات مستقبلية. ولتحقيق هذا الغرض لا بد من اللجوء إلى نظرية الانحدار التي تعطينا معادلة لرسم خط يشرح العلاقة بين Y,X ومنه نستطيع إجراء تنبؤات مستقبلية، ولتوضيح ذلك لا بد من التعرض إلى أمثلة:

1- من المعلوم أن للسماد الكيماوي دور في زيادة خصوبة التربة وزيادة قدرتها على إنتاج المحاصيل، لذلك فإن المحاصيل هي المتغير المعتمد Y والسماد هو المتغير الحر X . والجدول التالي يبين كميات الإنتاج المقابلة لكميات السماد المستعملة. والمطلوب توقيع أزواج القيم Y,X على ورقة رسم بياني لمعرفة تخطيط الانتشار ومنه سيتضح بأن هناك علاقة خطية بين المتغيرين ولكن النقاط لن تكون على خط مستقيم يسهل توصيله، لذا نلجأ إلى أسلوب الخط الممهد وهو أفضل خط يمكن رسمه بين هذه النقاط.

كمية السماد (قنطار/هكتار)	30	40	50	60	70	80
الإنتاج (قنطار/هكتار)	43	45	54	53	61	63

والمشكلة التي تواجهنا في هذا الأسلوب هو أن خط الانحدار الذي يرسمه كل منا بمفرده سوف يكون مخالفا لما يرسمه الآخرون نظرا لاختلاف حكمنا على أفضل خط يمكن أن يمثل هذه القيم، ولهذا أصبح ضروريا اللجوء إلى نظرية المربعات الصغرى التي تمكن من إيجاد ثوابت لمعادلة الانحدار b,a وهما أساس رسم خط الانحدار بصورة لا تختلف إطلاقا من شخص إلى آخر وليست عرضة لأهوائنا الشخصية.

طريقة المربعات الصغرى: Least Square Method

تكتب معادلة الخط المستقيم كما يلي:

$$Y=a+bX \qquad (12)$$

حيث أن: b,a ثابتان سبق شرحهما. والهدف من استخدام طريقة المربعات الصغرى هي تقدير قيمة كل من b,a لأجل استخدامهما في رسم الخط المستقيم.

إن طريقة المربعات الصغرى تعطينا أفضل توفيق للخط (Best fit) وفيها يتحدد خط الانحدار بحيث يكون مجموع مربعات انحرافات القيم المشاهدة(الفعلية) عنه أصغر ما يمكن. أي أن هذا المجموع يكون أقل من مجموع مربعات انحرافات هذه المشاهدات عن أي خط آخر. كما يتحدد خط الانحدار أيضا بحيث يكون مجموع انحرافات القيم عن الخط تساوي صفر، أي أن الانحرافات الموجبة التي تقع فوق الخط تواجهها انحرافات سالبة تقع أسفله. ويمكن حساب توابث معادلة الانحدار b,a بالمعادلتين التاليتين:

$$b=\frac{\sum XY-\frac{\sum X\sum Y}{n}}{\sum X^{2}-\frac{\left(\sum X\right)^{2}}{n}} \qquad (13)$$

$$a=\overline{Y}-b\overline{X} \qquad (14)$$

مثال1– استخدم بيانات الجدول التالي لحساب a و b لاستخدامهما في المعادلة رقم (3) لأجل رسم خط الانحدار.

جدول(27) العلاقة بين كمية السماد والإنتاج.

كمية السماد(قنطار/هكتار)	30	40	50	60	70	80
الإنتاج(قنطار/هكتار)	43	45	54	53	61	63

الحل:

أولاً:- نكون جدول لحساب عناصر معادلتي b,a كما يلي:

X	X^2	Y	YX
30	900	43	1290
40	1600	45	1800
50	2500	54	2700
60	3600	53	3180
70	4900	61	4270
80	6400	63	5040

$$\sum X = 330 \ldots\ldots\ldots\ldots \sum X^2 = 19900 \ldots\ldots\ldots\ldots \sum Y = 319 \ldots\ldots\ldots\ldots \sum XY = 18280$$

$$\overline{X} = \frac{330}{6} = 55 \ldots\ldots\ldots\ldots \overline{Y} = \frac{319}{6} = 53.17$$

ثانياً:- نعوض في المعادلة رقم (4) لحساب قيمة b وفي المعادلة رقم (5) لحساب قيمة a كما يلي:

$$b = \frac{18280 - \frac{330(319)}{6}}{19900 - \frac{(330)^2}{6}} = 0.42$$

$$a = 53.17 - 0.42(55) = 30.07$$

ثالثاً:- وبعد حصولنا على قيمتي b,a نستطيع استخدامهما في تقدير معادلة الخط المستقيم الخاصة بالعلاقة بين الإنتاج والسماد كما يلي:

Y=30.07+0.42X

ولرسم أي خط مستقيم نحتاج إلى معرفة نقطتين لتوصيل الخط بينهما. وبالنسبة لخط معادلة الانحدار لا تتوفر لنا إلا نقطة واحدة متمثلة في قيمة a التي تعبر عن قيمة Y عندما تكون قيمة X مساوية للصفر، لكن مع ذلك يمكن تقدير النقطة الثانية بأي من الطريقتين التاليتين:

1- عن طريق نقطة تقاطع $\overline{X}$ مع $\overline{Y}$ ، أي النقطة التي تتقاطع فيها 55 مع 53.17 .

2- عن طريق تعويض أي قيمة من قيم X القريبة من $\overline{X}$ في معادلة الخط المستقيم المقدرة للبيانات للحصول على قيمة Y. لنفرض أن X=60 فما هي قيمة Y التي تقابلها؟. نعوض في المعادلة كما يلي:

Y= 30.06+0.42(60) =55.26

وهكذا فإن النقطة التي نبحث عنها هي تقاطع (60، 55.26). والآن نرسم تخطيط الانتشار ونبين عليه خط الانحدار الذي تمثله المعادلة.

مثال2– تصنف الأمطار حسب الطرقة التي تم بها صعود الهواء وحدوث التكاثف في السحب. فهي إما أن تكون تصاعدية(انقلابية) نتيجة التسخين الشديد للهواء الرطب، وتتميز بسقوط رخات متقطعة وعنيفة خلال فترات قصيرة قد لا تزيد عن 15–20 دقيقة وعلى مساحات صغيرة. أي إنها أمطار محلية وشديدة العشوائية، فقد تسقط على قرية ولا تتعداها إلى قرية أخرى مجاورة لها.

والنوع الثاني من الأمطار هو التضاريسي الناتج عن مواجهة منحدرات جبلية مسار رياح رطبة، حيث يرغم الهواء على التراكم فوق بعضة ويضطرب ويصعد إلى أعلى في محاولته عبور الجبل إلى الناحية الأخرى، وأثناء صعوده يتكاتف بخار الماء مكونا السحب الكثيفة وتسقط منها الأمطار الغزيرة على السفوح المواجهة للرياح بينما تهبط هذه الرياح شبه جافه على السفوح الخلفية لذلك تسمى سفوح ظل المطر.

والنوع الثالث هو المطر الإعصاري الناتج عن تقابل كتلتين غير متجانستين من الهواء، حيث تتقدم الكتلة الباردة بإسفين من الهاء البارد والمضغوط تحت كتلة الهواء الدافئ مما يرغم الهواء الدفيء على الصعود وتتكاتف ما به من أبخرة مكونة السحب التي تتساقط منها الأمطار الغزيرة على طول جبهة الالتقاء لمدة طويلة ومغطية مساحات شاسعة من الكيلومترات المربعة.

التمرين التالي يحاول إثبات أن أغلب الأمطار الساقطة على منطقة الجفارة الليبية هي من النوع الثالث، أي أمطار الجبهات والأعاصير المتميزة بسقوطها على مساحات كبيرة، وهذا يعني تشابه كميات الأمطار الساقطة على محطات متقاربة. الجدول التالي يبين أمطار شهر نوفمبر بالسنتمترات لكل من محطتي أرصاد مطار طرابلس والعزيزية خلال الفترة 1957–1971. والمطلوب حساب معادلتي الارتباط والانحدار ، وقياس الدلالة الإحصائية لمعامل الارتباط عند معدل 0.95

جدول(28) العلاقة بين المطر في محطة المطار ومحطة العزيزية (ملم)

المطار	13.9	1.3	5.4	1.0	4.8	15.2	3.1	13.4	2.5	1.3	7.0	4.1	2.5	0.2	5.4
العزيزية	7.6	0.3	3.3	1.6	1.0	10.0	2.3	14.8	5.4	1.4	5.9	4.0	3.6	0.3	5.4

الحل:

أولاً:– نكون جدول لحساب عناصر معادلتي الانحدار والارتباط كما يلي:

المطار X	X^2	العزيزية Y	Y^2	YX
13.9	193.21	7.6	57.76	105.64
1.3	1.6	0.3	0.09	0.39
5.4	29.16	3.3	10.89	17.82
1.0	1.0	1.6	2.56	1.6
4.8	23.04	1.0	1.0	4.8
15.2	231.04	10.0	100.0	152.0
3.1	9.61	2.3	5.29	7.13
13.4	179.56	14.8	219.04	198.32
5.2	27.04	5.4	29.16	28.08
1.3	1.69	1.4	1.96	1.82
7.0	49.0	5.9	34.81	41.3
4.1	16.81	4.0	16.0	16.4
2.5	6.25	3.6	12.96	9.0
0.2	0.04	0.3	0.09	0.06
5.4	29.16	5.4	29.16	29.16

$$\sum X = 83.8 \ldots\ldots\ldots\ldots \sum X^2 = 798.3 \ldots\ldots\ldots\ldots \sum Y = 66.9 \ldots\ldots\ldots\ldots \sum Y^2 = 520.77$$
$$\sum XY = 613.52 \ldots\ldots\ldots\ldots \overline{X} = 5.58 \ldots\ldots\ldots\ldots \overline{Y} = 4.46$$

ثانياً:- نعوض في المعادلتين (4) و(5) لحساب ثابتي معادلة الانحدار:

$$b = \frac{613.52 - \frac{83.8(66.9)}{15}}{798.3 - \frac{(83.8)^2}{15}} = 0.73$$

$$a = 4.46 + 0.73(5.58) = 0.387$$

ثالثاً:- معادلة انحدار أمطار العزيزية على المطار هي:

$$Y = 0.387 + 0.73X$$

رابعاً:- معامل ارتباط أمطار العزيزية بالمطار هي:

$$r = \frac{613.52 - \frac{83.8(66.9)}{15}}{\sqrt{\left[798.3 - \frac{(83.8)^2}{15}\right]}\sqrt{\left[520.77 - \frac{(66.9)^2}{15}\right]}} = 0.88$$

خامساً:- اختبار الدلالة الإحصائية لمعامل الارتباط:

أ- الفرضيات:

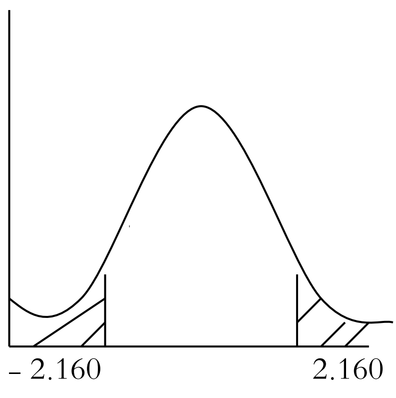
H0: r = 0

HA : r # 0

ب- الاختبار المناسب هو اختبار t :

$$t = 0.88\sqrt{\frac{15-2}{1-(0.88)^2}} = 6.6$$

ج- قيمة $t_{0.05}$ المقدرة عند 13 درجة حرية لاختبار الطرفين = + أو − 2.160 وبما أن الفرض البديل هو للطرفين نرسم المنحنى الثائي ونبين عليه موقع قيمتي t الجدولية لتحديد منطقة القبول في الوسط ومنطقتي الرفض في الطرفين، الأيمن الأكبر من 2.160 والأيسر الأصغر من − 2.160 كما يلي:



د- المقارنة والاستنتاج: بما أن t الإختبارية تساوي 6.6 أكبر من t الجدولية نرفض H0 ونستنتج بأن هناك ارتباط موجب دو دلالة إحصائية بين أمطار العزيزيه وأمطار المطار مما يعني بأنهما ينتميان إلى نظام الأعاصير وهو ما يؤيد الفرضية التي نصبوا إلى التأكد من صحتها.

تمرينات:

1- الجـدول التـالي لبيانـات المطـر وعـدد الأيـام الممطـرة بالزنتـان خـلال الفـترة 1980-1993 .احسب معادلة انحدار المطر على عدد الأيام الممطرة.

جدول(29) العلاقة بين المطر السنوي وعدد الأيام الممطرة في السنة بمحطة أرصاد مدينة الزنتان.

السنة	كمية المطر	عدد الأيام الممطرة
1980	472.1	32
1981	98.6	26
1982	284.1	44
1983	254.7	35
1984	347.2	31
1985	294.3	31
1986	363.9	41
1987	157.3	24
1988	209.5	26
1989	234.5	22
1990	87.1	18
1091	357.0	49
1992	93.8	20
1993	205.0	31

2- الجدول التالي يبين قياسات المطر السنوي X ومعدل انجراف التربة من 11 نقطة، والمطلوب ما يلي:

أ- استخدام هذه البيانات في حساب معادلة انحدار Yعلى X.

ب- رسم تخطيط الانتشار مبينا عليه معادلة الخط المستقيم.

ج- اختبار الدلالة الإحصائية لمعامل الارتباط، $\alpha_{0.05}$.

جدول (30) العلاقة بين كمية المطر والانجراف(بآلاف المليمترات)

المطر X	71	72	73	77	70	87	91	95	103	105	11.2
الانجراف Y	5.7	8.2	9.4	9.6	9.8	10.2	11.9	11.7	13.4	13.2	13.4

لاحظ أن قيم Y بآلاف المليمترات/ بالمتر المكعب/ ساعة. قيم X آلاف المليمترات.

وللحصول على معادلة تقديرية بالوحدات القياسية الصحيحة يجب ضرب a في 1000، أما المعامل b فلا يتأثر بتغير المقياس.

3- عند التخطيط لمشاريع السدود والطاقة الكهرومائية فإنه من المهم جدا معرفة تاريخ الجريان النهري، ولكن في بعض الأحيان يصعب الحصول على هذه البيانات مما يضطر الباحثين إلى تقديرها من بيانات محطة أخرى قريبة. ففي منتصف الأربعينات من القرن الماضي واجهت الباحثين مشكلة قصر سجل جريان نهر كوتيناي Kootenai بإقليم كولومبيا البريطانية حيث كانوا يخططون لإنشاء سد مائي، فلم تزد مدة السجل حينذاك عن سنة 1931، ولكن توجد في الوقت نفسه سجلات أطول على مسافة بعيدة من الموقع عند أدنى المجرى على الحدود الأمريكية بمنطقة ليبيLibby بولاية مونتانا. فهل بالإمكان تقدير كميات جريان كوتيناي بإقليم كولومبيا من بيانات منطقة ليبيLibby على الحدود الأمريكية.

جدول(31) الجريان السنوي عند نيجاويت وليبي

جريان نيجاويت Y	جريان ليبي X	السنة
19.7	27.1	1931
18.0	20.9	32
26.1	33.4	33
44.9	77.6	34
26.1	37.0	35
19.9	76.6	36
15.7	17.6	37
27.6	35.1	38
24.9	32.6	39
23.4	26.0	1940
23.1	27.6	41
31.3	38.7	42
23.8	27.8	1943

والمطلوب:

أ- رسم تخطيط الانتشار

ب- حساب معادلة انحدار Y علىX ومعامل الارتباط r

ج- اختبار الدلالة الإحصائية لمعامل الارتباط، استخدم مستوى 0.95

د- احسب كمية الجريان عند نيجاويت إلى سنة 1925 إذا كان الجريان عند ليبي لنفس الفترة كما يلي:

جدول(32) الجريان السنوي عند نيجاويت للفترة 1925–1930.

السنة X	1925	1926	1927	1928	1929	1930
الجريان Y	42.0	24.0	38.0	49.4	24.6	24.2

4- الجدول التالي يبين تطور عدد المشتغلين بالزراعة في أمريكا(ملايين) خلال الفترة 1949-1957.

جدول(33) تطور عدد المزارعين بالولايات المتحدة خلال الفترة 1949–1957.

السنة X	1	2	3	4	5	6	7	8	9
عدد المزارعينY	9.96	9.93	9.55	9.15	8.86	8.64	8.36	7.82	7.58

Murray,R.Spiegle.Statistics.1961.p.239.

المطلوب:

أ- تقدير معادلة انحدار y على X ورسم خط الانحدار على ورقة تخطيط الانتشار.

ب- تقدير عدد المزارعين لسنة 1948 ومقارنته بالرقم الفعلي الذي هو 10.36

ج- تقدير عدد المزارعين لعام 1958 ومقارنته بالرقم الفعلي الذي هو 7.53 ، هل المعادلة تبالغ التقدير أو تبخس التقدير.

د- وضع مجموعة من الافتراضات لشرح هذا التناقص المستمر في أعداد المزارعين.

5- الجدول (34) يبين تطور أعداد الأطباء الوطنيين وأعداد المرضى لكل طبيب في ليبيا خلال الفترة 1969-1978، **والمطلوب:**

أ- حساب معادلة انحدار عدد الأطباء على الزمن.

ب- حساب معادلة انحدار عدد المرضى على الزمن.

ج- حساب معامل الارتباط بين عدد الأطباء والمرضى.

جدول(34) العلاقة بين عدد الأطباء الليبيين وعدد المرضى الين يخدمونهم.

السنة	1	2	3	4	5	6	7	8	9	10
عدد الأطباء X	733	731	1336	1525	1736	2063	2585	2903	2926	3350
مريض لكل طبيب Y	2588	2707	1547	1416	1300	1139	916	892	822	809

المصدر: ثورة الفاتح في عشر سنوات، دار الشورى، بيروت، 1979، ص.62.

6- فيما يلي جدول يبين تطور إحصاءات السكان بكندا من سنة 1871 إلى 1991. احسب معادلة الانحدار التنبؤية ومعامل الارتباط ثم قدر عدد السكان لعام2003 وقارنه بالفعلي الذي كان32.207

جدول(35) تطور عدد السكان بكندا خلال الفترة 1871– 1991.

عدد السكان (مليون)Y	السنةX
3.689	1871
4.833	1891
7.206	1911
10.376	1931
14.009	1951
21.568	1971
27.296	1981

7- الجدول التالي يبين تطور الحمولة الكهربائية لشبكات الجماهيرية من سنة 1990 إلى سنة 2003، احسب معامل ارتباط بيرسون ومعادلة الانحدار ثم قدر الحمولة لسنة 2010.

جدول(36) تطور الحمولة الكهربائية في ليبيا من سنة 1990 إلى 2003.

الحمولة **Y**	السنة **X**
1634	1990
1715	1991
1796	1992
1776	1993
1857	1994
2038	1995
2119	1996
2200	1997
2281	1998
2361	1999
2443	2000
2523	2001
2604	2002
2685	2003

المصدر: مجلة الطاقة والحياة، العدد(3) يونيو 1994، ص45 .

مما تقدم يتبين أن نموذج الانحدار البسيط يحاول شرح التغيرات التي تحصل في المتغير التابع Y الناتجة عن تغيرات X، أو بمعنى آخر، كيف والى أي مدى تعتمد التغيرات في المتغير التابع على تغيرات في المتغير الحر.

اختبار الدلالة الإحصائية:

توجد العديد من اختبارات الدلالة تؤدي جميعها إلى نفس النتيجة وهي رفض أو قبول الفرضية الصفرية اخترنا منها الاختبارين التاليين:

أ– اختبار نسبة التباين التي يشرحها r^2 :

معامل التحديد r^2 يعني نسبة التباين في Y التي يشرحها X أما $\left(1-r^2\right)$ فهي تعني نسبة التباين في Y الناتجة عن الأخطاء. وتحسب قيمة r^2 بمربع قيمة معامل ارتباط بيرسون المحسوب بالمعادلة رقم (1) التي سبق شرحها وهي:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\left[\sum X^2 - \frac{\left(\sum X\right)^2}{n}\right]}\sqrt{\left[\sum y^2 - \frac{\left(\sum y\right)^2}{n}\right]}} \quad \text{..............................(15)}$$

خطوات إجراء الاختبار:

1- الفرضيات:

2- الاختبار الإحصائي:

أولاً:- الفرضيات:

H0: $r^2 = 0$

HA: $r^2 \# 0$

ثانياً:- الاختبار:

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

ثالثاً:- وتقارن قيمة t الاختبارية مع t_α المقدرة لاختبار الطرفين بدرجات الحرية (n-2) ، ونرفض H0 إذا جاءت t في الإقليم الحرج.

ب- اختبار دلالة معامل الانحدار b:

لاختبار معامل الانحدار نتبع الخطوات التالية:

1- الفرضيات:

H0 : b = 0

HA : b # 0

2- حساب قيمة الخطأ المعياري Se :

$$Se = \sqrt{\frac{\left(\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right) - b}{n-2}}$$

والخطأ المعياري عبارة عن مقياس لتشتت الأخطاء ويعني درجة انتشار النقاط حول خط الانحدار.

3- حساب الانحراف المعياري لمعامل الانحدار (Sb):

$$Sb = \frac{Se}{\sqrt{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}}$$

4- حساب قيمة الاختبار الإحصائي(t):

$$t = \frac{b}{Sb}$$

5- تقدير قيمة $t_{(\alpha,n-2)}$ لاختبار الطرفين والمقارنة والاستنتاج:- نرفض H0 إذا جاءت t بالإقليم الحرج مما يعني أن معامل الانحدار مختلف جذريا عن الصفر، وبعبارة أخري التغيرات الحاصلة في Y ناتجة عن تغيرات X.

أمثلة:

1- المعطيات التالية محسوبة لبيانات الزراعة الأمريكية، اختبر الدلالة الإحصائية لمعامل الانحدار (b) باستخدام مستوى ثقة 0.95.

$$n = 9..........\sum X = 45...........\sum X^2 = 285.........\sum Y = 79.85........\sum Y^2 = 714.38$$

$$\sum XY = 380.5.............r^2 = 0.98...............a = 10.45.............b = -0.31$$

الحل:

معادلة الانحدار البسيط للزراعة الأمريكية هي:

Y=10.45-0.31X

أولاً:- الفرضيات:

H0 : b = 0

HA : b # 0

ثانياً: – حساب الخطأ المعياري، نعوض عن المعطيات في معادلة الخطأ المعياري:

$$Se = \sqrt{\frac{\left[714.38 - \frac{(79.85)^2}{9}\right] - (-0.31)}{9-2}} = 0.9443$$

ثالثاً:- حساب الانحراف المعياري لمعامل الانحدار:

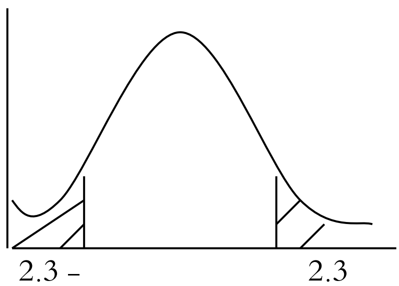
$$Sb = \frac{0.9443}{\sqrt{285 - \frac{(45)^2}{9}}} = 0.1219$$

رابعاً:- حساب قيمة الاختبار: تحصلنا على قيمتي الخطأ المعياري Se(0.9443) والانحراف المعياري Sb (0.1219) نعوض في المعادلة ونحسب قيمة الاختبار الثائي كما يلي:

$$t = \frac{0.9443}{0.1219} = 7.7$$

خامساً:- قيمة $t_{(0.05,7)}$ لاختبار الطرفين = +2.3 أو – 2.3 (أنظر الشكل التالي)

وبما أن قيمة t = 7.7 واقعة بالإقليم الحرج نرفض H0 ونستنتج بأن هناك تأثير أكيد تفرضه Xعلى Y.



تمرينات إضافية:

1- استخدم بيانات الجدول التالي لإيجاد الارتباط ومعادلة الانحدار البسيط بين معدل انجراف التربة وشدة المطر.

الحل:

جدول(37) العلاقة بين شدة المطر ومعدل الانجراف بعدد 11 محطة قياس.

شدة المطر	7.1	7.2	7.3	7.7	7.6	8.7	9.1	9.5	10.3	10.5	11.2
معدل الانجراف	5.7	8.2	9.4	9.6	9.8	10.2	11.9	11.7	13.4	13.2	13.4

وتأكد من أن نتائج التحليل هي كما يلي:

$R=0.913$

$R^2=0.933$

$a=-2.469$

$b=1.493$

$Y=-2.469+1.493X$

المطلوب: اختبار الدلالة الاحصائية لمعامل الانحدار b

2- البيانات التالية تخص العلاقة بين البعد من مركز المدينة وإيجار الشقق، أوجد علاقة الارتباط ومعادلة الانحدار الخطي التي تشرح هذه العلاقة وتتنبأ بها.

جدول (38) العلاقة بين المسافة من مركز المدينة وإيجار الشقق.

المسافة كم	4	10	17	19	31	34	36	41	49	58
الإيجار	97	35	88	70	63	36	42	62	35	41

تأكد من أن نتائج التحليل كما يلي:

$r=0.599$

$r^2=0.359$

$a=80.56$

$b=-0.792$

$T=6.3$

Sig at 0.00

$Y=80.56-0.792X$

3- الجدول التالي يبين حالات الملاريا السنوية وكميات الأمطار المسجلة بزمبابوي. عند مستوى دلالة إحصائية 0.95 اختبر هل هناك علاقة ارتباط بين تفشى الملاريا وكميات الأمطار(ملم).

جدول(39) حالات الملاريا وكميات المطر السنوية بزمبابوي.

السنة	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
المطر	450	750	600	620	500	330	630	500	400	700	800
الملاريا	42000	96000	70000	60000	58000	40000	81000	78000	72000	17000	188000

4- الجدول التالي لمصروفات ميزانية التحول حسب البلديات للفترة 1981-1989.

احسب ارتباط بيرسون بين عدد السكان ونسبة المصروفات ثم اختبر الدلالة الإحصائية لمعامل الارتباط، مستوى ثقة 0.95

جدول(40) مصروفات ميزانية التحول حسب البلديات.

البلدية	عدد السكان عام 1989 (الف نسمة)	نسبة المصروفات %
البطنان	119.5	2.6
الجبل الأخضر	413.5	10.3
بنغازي	611.1	15.2
خليج سرت	491.6	12.8
المرقب	482.4	10.3
الجبل الغربي	243.5	7.2
طرابلس	1249.7	18.0
الزاويه	385.9	5.6
النقاط الخمس	294.1	5.9
سبها	156.2	4.5
مرزق	55.0	2.2
وادي الحياه	59.7	4.2
الكفره	32.2	1.2
المجموع	**4594.4**	**100.0**

5- البيانات التالية للمعدلات الشهرية للمطر وعدد الأيام الممطرة بمنطقة الزاوية والمطلوب حساب معامل الارتباط بين كميات المطر وعدد الأيام الممطرة.

جدول (41) المعدلات الشهرية للمطر وعدد الأيام الممطرة بمدينة الزاوية.

الشهر	1	2	3	4	5	6	7	8	9	10	11	12
المطر(ملم)	55.3	38.2	29.2	9.9	6.5	1.3	0.1	0	12.4	24.1	59.1	55.9
الأيام الممطرة	8.5	5.8	4.2	2.8	1.9	0.5	1	0	2.5	4.3	6.3	7.9

6- استخدم بيانات الجدول التالي في حساب معادلة الانحدار بين الزمن وعدد وفيات التورنيدو المسجلة بالولايات المتحدة خلال الفترة 1916-1997.

جدول(42) الوفيات السنوية لأعاصير التورنيدو بالولايات المتحدة المسجلة للفترة الممتدة من 1916 حتى 1977. لاحظ أن الجدول يسير من اليمين إلى اليسار بداية من عام 1916 وقيمته 150 وانتهاء بعام 1977 وقيمته 67.

150	551	136	206	499	202	135	110	376	794	144	540
95	274	179	36	394	362	47	71	552	29	183	91
65	53	334	58	275	210	78	313	139	211	70	34
229	515	36	126	83	192	66	58	46	51	28	31
73	296	98	114	131	66	72	156	27	87	361	60
44	43	53	84	28	24	64	34	122	94	15	59
32	50	53	39	39	33	69	30	25	67		

المصدر: باتريك . ل . أبوت، ترجمة توفيق علي منصور، **الكوارث الطبيعية**، المجلد الثاني، المجلس الأعلى للثقافة، القاهرة، 2003.

7- الجدول التالي يبين الوفيات وخسائر الممتلكات جراء الحرائق المسجلة بالولايات المتحدة خلال الفترة 1978- 1994. والمطلوب حساب معادلة الارتباط والانحدار بين الزمن والخسائر ورسم معادلة الانحدار.

جدول(43) الوفيات وخسائر الممتلكات جراء الحرائق المسجلة بالولايات المتحدة خلال الفترة 1978- 1994.

السنة	الوفيات	الأضرار بلايين الدولارات
1978	8130	4008
1979	7780	4851
1980	6505	5579
1981	6700	5625
1982	6020	5894
1983	5920	6320
1984	5240	7602
1985	6185	7753
1986	5850	8488
1987	5810	8504
1988	6215	9626
1989	5410	9514
1990	5195	9495

المصدر: باتريك . ل . أبوت، ترجمة توفيق علي منصور، **الكوارث الطبيعية**، المجلد الثاني، المجلس الأعلى للثقافة، القاهرة، 2003.

8- ارجع إلى بيانات المطر لكل من مطار طرابلس والعزيزية واحسب معادلتي الانحدار والارتباط ثم اختبر الدلالة عند مستوى 0.95 لكل من معامل الارتباط التربيعي ومعادلة الانحدار.

9- الجدول التالي لإنتاجية أحد المصانع والمطلوب إيجاد ما يلي:

أولاً:- معادلة انحدار كل من:

أ- عدد القطع والوزن.

ب- عدد القطع ومدة التصنيع.

ج- الوزن ومدة التصنيع.

ثانياً:- معامل الارتباط التربيعي لكل من الفقرات أ، ب، ج.

ثالثاً:- تخطيط الانتشار مبينا عليه معادلة خط الانحدار لكل من الفقرات أ، ب،ج.

رابعاً:-كم سيكون الوزن إذا زاد عدد القطع إلى 70.

خامساً:-كم مدة التصنيع عندما يصل الوزن إلى 80 قنطارا.

جدول(44) إنتاجية أحد المصانع.

الشحنة	عدد القطع	الوزن(قنطار)	مدة التصنيع(دقيقة)
1	55	60	60
2	20	25	8
3	35	40	17
4	45	50	23
5	40	45	19
6	25	30	11
7	55	60	28
8	30	35	15
9	60	65	32
10	45	50	23
11	35	40	17
12	25	30	11
13	45	50	23
14	35	40	17
15	30	35	14

John Neter & Others. Applied Statistics. 1978. P.513.

10- قام الباحثان كريستودولو ولويس سنة 1976 بأخذ عينة متكونة من 40 طالبا لأجل تقدير المسافات الفاصلة بين 20 جوز من النقاط على الأرض ثم حسبت قيم Y على أساس متوسط تقديرات الطلبة للمسافة بين كل نقطتين. أما قيم X فهي تمثل المسافة الفعلية. الجدول التالي يبين المسافات الفعلية والمقدرة، والمطلوب حساب معادلة الانحدار ثم اختبار الدلالة عند مستوى 0.95 لمعامل الانحدار b وأخيرا رسم تخطيط الانتشار مبينا عليه معادلة خط الانحدار.

جدول(45) المسافة المقدرة والحقيقية ل20 جوز من النقاط بالكيلومترات.

المقدرة Y	الفعلية X	المقدرة Y	الفعلية X
0.8	0.9	2.0	1.3
1.2	0.6	0.6	0.6
0.5	0.5	1.0	0.9
2.6	1.9	1.3	1.2
0.7	0.7	1.8	1.2
0.9	0.8	3.1	2.9
2.4	2.2	1.1	0.7
1.8	1.6	1.6	0.8
1.2	1.2	2.4	2.3
1.2	1.0	0.9	0.6

Christodolou, F. and A. d. Lewis. Distance Perception Of Reading Univ. Students. Univ. of Reading. Dept of Geography.1976.

11- الجدول التالي يبين تطور أعداد السكان وجرائم القتل بالولايات المتحدة الأمريكية والمطلوب ما يلي:

أ- رسم تخطيط الانتشار مبينا عليه معادلة خط الانحدار بين جرائم القتل وأعداد السكان.

ب- اختبار الدلالة الإحصائية لمعامل الانحدار عند مستوى 0.95

ج- حساب معامل الارتباط التربيعي بين الزمن وجرائم القتل واختبار دلالته الإحصائية عند مستوى 0.95

الجدول(46) تطور أعداد السكان وجرائم القتل بالولايات المتحدة الأمريكية خلال الفترة 1964-1973م.

السنة	عدد السكان(مليون)	عدد جرائم القتل
1964	191.9	4393
1965	194.3	5015
1966	196.6	5660
1967	198.7	6998
1968	200.7	8105
1969	202.7	8876
1970	204.9	9039
1971	207.1	10712
1972	208.8	10379
1973	210.4	11249

المصدر: باتريك . ل . أبوت، ترجمة توفيق علي منصور، **الكوارث الطبيعية**، المجلد الثاني، المجلس الأعلى للثقافة، القاهرة، 2003.

12- الجدول التالي يبين معدل الإنتاج السنوي لمحصول القمح بولاية نيويورك الأمريكية للعقود المنتهية أدناه. احسب معادلة الانحدار واختبر الدلالة الإحصائية لمعامل الانحدار عند مستوى 0.01 .ثم تنبأ بكم سيكون عليه الإنتاج في عقد 1985 وقارنه بالفعلي الذي كان(43.3 بوشل \ أيكر).

جدول(47) معدل الإنتاج السنوي لمحصول القمح بولاية نيويورك الأمريكية للعقود المنتهية أدناه.

العقد	الإنتاج
1875	16.1
1885	17.5
1895	18.0
1905	17.9
1915	20.1
1925	19.3
1935	19.2
1945	23.9
1955	27.9
1965	33.5
1975	39.3

Neal,F.Jensen:Limits to Growth in World food Production. Science.Vol (201).28 July 1978.P.317.

13- فيما يلي كمية المطر السنوي (ملم) وعدد الأيام الممطرة بمدينة نالوت خلال الفترة 1968-1988. والمطلوب حساب معادلة انحدار المطر على عدد الأيام الممطرة ثم التنبؤ بكمية المطر إذا كانت عدد الأيام الممطرة 50.

جدول(48) كمية المطر السنوي(ملم) وعدد الأيام الممطرة بمدينة نالوت
خلال الفترة 1968-1988

عدد الأيام الممطرة	كمية المطر(ملم)	عدد الأيام الممطرة	كمية المطر(ملم)
17	57.5	27	218.1
36	130.2	28	165.1
16	97.0	19	35.9
18	78.9	38	225.3
38	291.9	27	145.2
24	55.7	28	136.1
31	205.5	17	183.2
32	148.4	26	187.2
46	568.8	23	157.2
19	100.6	24	153.4
18	75.0	-	-

الفصل الخامس

الاختبارات اللامعلمية

الاختبارات اللامعلمية **Non-Parametric Tests** تعني أنه بإمكاننا إجراء الاختبارات الإحصائية بدون الحاجة إلى معالم العينة المتمثلة في المتوسط الحسابي والانحراف المعياري.

1- اختبار الاستقلالية:

ويسمى أحيانا اختبار **مجموع الرتب** ويسمى أحيانا أخري **اختبار U لمانوتني**: ويستخدم هذا الاختبار لمعرفة هل العينتان أخذتا باستقلالية تامة من مجتمعين متماثلين في الوقت الذي يوجد فيه ما يوحي بأن شروط استخدام الاختبار التائي T غير متوفرة. ولتسهيل الموضوع للفهم نورد المثال التالي:

الجدول التالي يبين تبرعات مجموعتين من ربات البيوت أخذتا عشوائيا من منطقتين سكنيتين يعتقد أنهما من نفس المستوى الاقتصادي، والمطلوب اختبار فرضية تساوي التبرعات في المنطقتين A و B.

المنطقةA	120	155	88	420	360	60	650	82	–	–	–	–
المنطقةB	42	190	320	680	82	130	500	32	160	82	890	1020

للوهلة الأولى ربما نفكر في مقارنة المتوسط الحسابي لكل منهما باستخدام الاختبار التائي الذي سبق شرحه، لكن بتفحص الجدول مليئا يتبين أن هناك تغيرات كبيرة(تباين) في تبرعات المجموعة B أكبر مما هي في تبرعات المجموعة A مما يلقي ظلالا من الشك على فرضية تساوي التباين، أضف إلى ذلك أن فرضية التوزيع ألاعتدالي للبيانات هي الأخرى مفتوحة للتساؤل. وبالمناسبة يجب أن نتذكر أنه لاستخدام الاختبارات المعلمية Z و T التي سبق شرحها دائما نفترض أن يكون توزيع المشاهدات اعتداليا والتباين متساوي على جميع مراحل العينة. لكن إذا كانت التوزيعات المدروسة لا تلبي هذين الشرطين فإن استخدام الاختبار الزائي والتائي سوف يقود إلى نتائج مغالطة للواقع.

في هذا الاختبار تدمج مفردات العينتين وكأنهما من عينة واحدة أكبر، ثم ترتب المفردات تصاعديا بحيث تتحصل أصغر القيم على الترتيب رقم 1، أما التي تليها فتعطى الترتيب رقم 2 وهلم جرا حتى نكمل جميع المفردات كما في الجدول التالي:

جدول (49) تبرعات المنطقة A وB مجمعة ومرتبة تصاعديا

الترتيب	المنطقة	التبرعات
1	b	32
2	b	42
3	a	60
5	b	82
5	b	82
5	b	82
7	a	88
8	a	120
9	b	130
10	a	155
11	b	160
12	b	190
13	b	320
14	a	360
15	a	420
16	b	500
17	a	650
18	b	680
19	b	890
20	b	1020

لاحظ أنه عندما تتساوى قيم مجموعة من المفردات، تعطى متوسط رتبها الحسابي، فعلى سبيل المثال، المفردة الرابعة والخامسة والسادسة تتساوى في القيمة 82 عليه فإن رتب هذه المفردات يساوي (4+5+6)÷ 3 =5، وكنتيجة لهذا الترتيب فإن رتب المجموعة A تأخذ النمط التالي:

3 5 7 8 10 14 15 17

أما رتب المجموعة B فتأخذ النمط التالي:

20 19 18 16 13 12 11 9 5 5 2 1

لذلك فإن مجموع رتب A = 79

ومجموع رتب B = 131

فرضية العدم H0 المراد اختبارها تقول بأن تبرعات كلا المجموعتين تتصف بنفس التوزيع.

الفرضية البديلة HA تقول بوجود اختلاف جوهري في توزيعيهما.

لنفرض أن m = حجم العينة الأولى.

N = حجم العينة الثانية.

R1 = مجموع رتب العينة الأولى.

R2 = مجموع رتب العينة الثانية.

الاختبار الإحصائي:

$$U = MN + \left[\frac{M(M+1)}{2}\right] - R1$$

ويحسب المتوسط الحسابي للرتب بالمعادلة التالية:

$$\overline{X}u = \frac{MN}{2}.$$

أما الانحراف المعياري للرتب فيحسب بالمعادلة التالية:

$$Su = \sqrt{\frac{MN(M+N+1)}{12}}$$

وإذا كانت M و N أكبر من 8 فإن توزيع U يأخذ الشكل ألاعتدالي تقريبا ويصير بالإمكان استخدام الاختبار الزائي كما يلي:

$$Z = \frac{U - \overline{X}u}{Su}$$

مثال: استخدم بيانات الجدول السابق لحساب U لاستخدامها في معادلة Z لاختبار فرضية تساوي تبرعات المجموعتين.

أولاً:- المعطيات والفرضيات:

M=8

N=12

R1=79

H0 : تبرعات كلا المجموعتين تتصف بنفس التوزيع.

HA : هنالك اختلاف جدري في توزيع تبرعات المجموعتين.

$$U = 8(12) + \left[\frac{8(8+1)}{2}\right] - 79 = 53$$

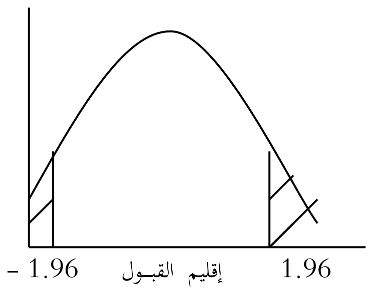
$$\overline{X}u = \frac{8(12)}{2} = 48$$

$$Su = \sqrt{\frac{8(12)(8+12+1)}{12}} = 12.96$$

ثانياً:- الاختبار: نعوض المعطيات في المعادلة الاختبارية :

$$Z = \frac{53-48}{12.96} = 0.39$$

ثالثاً:- تحديد المنطقة الحرجة والمقارنة والاستنتاج: بما أن الفرضية البديلة تقول بوجود اختلاف فهي لاختبار الطرفين لذلك فقيمة Z عند مستوي 95%=1.96 و- 1.96 .



وبما أن Z واقعة ضمن حدود الثقة نقبل H0 ونستنتج بأنه ليس هناك ما يشير إلى وجود اختلاف جوهري في توزيع تبرعات المنطقتين.

مثال 2- فصل متكون من 18 طالبا قسم إلى مجموعتين، الأولى تتكون من 8 والثانية من 10 طلاب. أعطيت الأولى تدريبات خاصة استخدمت خلالها وسائل سمعية وبصرية مساعدة، أما المجموعة الثانية فاقتصرت على محاضرات نظرية فقط. ثم أجري نفس الامتحان على المجموعتين في

وقت واحد، وسجلت درجاتهم في الجدول التالي، أوجد هل أدى اختلاف طريقة التدريس إلى حصول اختلاف جدري في نتائج المجموعتين، استخدم نسبة ثقة 95%.

A	95	66	42	82	71	68	52	86	–	–
B	49	39	80	65	30	62	53	49	55	60

الحل:

أولاً:– الفرضيات:

H0 : كلا الطريقتان تؤديان إلى نفس النتائج.

HA : الطريقتان تؤديان إلى نتائج مختلفة.

ثانياً:– تكوين جدول يجمع بين مفردات المجموعتين وترتب فيه المفردات تصاعديا كما لو كانت من مجتمع واحد مع إعطاء كل مفردة صنف المجموعة التابعة لها.

الرتبة	المجموعة	الدرجة
1	B	30
2	B	39
3	A	42
4.5	B	49
4.5	B	49
6	A	52
7	B	53
8	B	55
9	B	60
10	B	62
11	B	65
12	A	66
13	A	68
14	A	71
15	B	80
16	A	82
17	A	86
18	A	95

ثالثاً:- المعطيات الجاهزة والمحسوبة:

R1=3+6+12+13+14+16+17+18=99

M=8

N=10

$$U = 8(10) + \left[\frac{8(8+1)}{2}\right] - 99 = 17$$

$$\overline{X}u = \frac{8(10)}{2} = 40$$

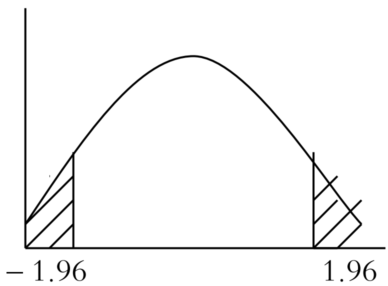
$$Su = \sqrt{\frac{8(10)(8+10+1)}{12}} = 11.25$$

رابعاً:- الاختبار:

$$Z = \frac{17-40}{11.25} = -2.04$$

خامساً:- تحديد الإقليم الحرج والمقارنة والاستنتاج:

بما أن الفرضية البديلة هي لاختبار الطرفين فإن قيمة Z_α=1.96 أو − 1.96



وبما أن Z جاءت بالإقليم الحرج نرفض H0 ونستنتج بأن هنالك اختلاف جوهري في نتائج الأسلوبين بسبب اختلاف الوسيلة التعليمية.

تمرينات:

1- لقد تم اختيار 10 منتخبين عشوائيا من بين أولئك المؤيدين لمشروع أحد القرارات، وتم أيضا اختيار 12 منتخبا من المصوتين ضده. الجدول التالي يبين أعمار أفراد العينتين، استخدم اختبار مجموع الرتب لاختبار الفرضية القائلة بتساوي توزيع أعمار الفئتين المؤيدة والمعارضة ضد الفرضية البديلة القائلة بوجود اختلاف بينها. مستوى الثقة 0.95 .

مؤيدين	28	33	27	31	29	25	58	30	25	41	–	–
غير مؤيدين	31	43	45	37	40	41	48	35	33	39	42	36

2- البيانات التالية لعدد الدقائق التي يتطلبها مجموعتين من العمال لتعلم بعض المهارات، استخدم اختبار مجموع الرتب لاختبار H0 القائلة بأن العينتان مأخوذتان من مجتمع واحد(أي لا يوجد اختلاف في التوزيع) ضد HA القائلة بوجود اختلاف. α=0.05

A	72	66	77	69	74	80	65	81	75	71
B	74	79	62	82	78	86	76	71	70	67

3- في نهاية العام الدراسي تم اختبار 12 طالبا نظاميا و 15 طالبا منتسبا في مادة المناخ ودونت نتائجهم في الجدول التالي، استخدم α 0.05 لاختبار هل المجموعتين من نفس المجتمع، أم أن هناك اختلاف جوهري بينهما يدل على أنهما من مجتمعين مختلفين.

نظامي	46	29	85	61	87	35	57	87	76	79	49	92	–	–	–
منتسب	92	41	36	72	34	67	99	72	83	82	76	40	84	67	24

4- أجري سباق بين نوعين من الخيول ورتبت نتائجها تصاعديا في الجدول التالي، هل صحيح بأنه لا يوجد اختلاف في توزيع رتب المجموعتين. α 0.05

النوع	a	b	b	a	a	a	a	b	a	a	a	b	a	b	b	a	b	b	b
الرتب	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

5- لمعرفة تأثير هرمون معين في الحد من نمو براعم النباتات، أخذت عينة عشوائية متكونة من 22 نبتة وقسمت إلى مجموعتين 11 لكل منهما، والمجموعة الأولى عولجت بهرمون وتركت الثانية بدون معالجة، وبعد أسبوعين دونت أطوال البراعم (مليمترات) في الجدول التالي. اختبر الفرضية القائلة بأن توزيع أطوال النبتات متساوي ولا أثر للهرمون في النمو ضد الفرضية البديلة HA. α 0.05

معالجة	26	28	37	41	57	58	67	78	80	121	122
غير معالجة	15	30	30	65	91	103	132	135	137	141	198

2- اختبار المتلاحقات Runs Test

ويسمى أيضا اختبار العشوائية Randomness أو الاعتدالية Normality في توزيع المشاهدات. إذ أن جميع الاختبارات اللإحصائية المعلمية بنيت على افتراض أن مشاهدات العينات المستخدمة في التحليل موزعة عشوائيا وليست منحازة إلى فئة دون أخرى، غير أن هناك حالات يصعب فيها التحقق من هذا الافتراض لذلك يجب اختبار عشوائية العينة قبل التعرض إلى التحليل وإصدار الأحكام بقبول أو رفض الفرضيات. وتظهر هذه الحالات بصفة خاصة عندما نكون عاجزين كليا أو جزئيا عن التحكم في اختيار العينة. فمثلا، في تقدير معدل الوفيات من مرض معين لا مفر من الاعتماد على سجلات سابقة وهذه لا تشكل عينة عشوائية بالمعنى الدقيق، وكذلك الحال لا يكون لنا خيار إلا الاعتماد على سجلات متاحة لإجراء تنبؤات جوية عن أحوال الطقس أو لدراسة حوادث المرور أو حين نأخذ وحدات صناعية بحسب ترتيب إنتاجها.

إن نظرية المتلاحقات تعتمد على الترتيب ألذي سحبت بها مشاهدات العينة. ويهدف الاختبار إلى كشف عما إذا كان هذا الترتيب عشوائيا أو يتخذ نمطا معينا لا يمكن أن نعزوه إلى الصدفة. فإذا كانت المشاهدات عشوائية فإنها ستتمشى مع التوزيع الاعتدالي، أما إذا أظهرت اتجاها أو تغيرا دوريا نتيجة لمؤثرات طبيعية أو بشرية فإنه لا يمكن استخدام اللاختبارات المعلمية لأنها تقود إلى استنتاجات خاطئة حيث سنرفض H0 في الوقت الذي يجب أن نقبلها فيه أو العكس. إن اختبار المتلاحقات يكشف لنا فيما إذا كان توزيع المشاهدات عشوائيا أم يظهر اتجاها أو تبدلا دوريا. إن هذا الاختبار حساس ولا يتأثر بشكل التوزيع التكراري للمشاهدات، فإذا كانت عشوائية وهذا ما هو متوقع منها فإنه يمكن استخدام الاختبارات المعلمية التي سبق شرحها، أما إذا أظهرت غير ذلك فيجب البحث عن السبب وراء ذلك.

تأمل النسق الذي تتخذه المتلاحقات التالية في الرمزين d و r المفصولين بخط مائل (/).

dddd/rrr/d/rr/ddd/rr

هنالك 6 متلاحقات منهم 3 تابعة للرمز r و3 تابعة للرمز d. إن القليل من المتلاحقات دليل على قلة العشوائية. افرض أنه تمت مقابلة 25 شخصا لتحديد نوع انتمائهم السياسي بحيث يرمز للديمقراطي بالرمز d والجمهوري بالرمز r وجاءت إجاباتهم كما يلي:

ddd/rr/d/rr/ddd/rrrr/ddd/r/ddd/r/dd

افرض أن: r = عدد المتلاحقات=11

m = عدد رموز d =15

N = عدد رموز r =10

خطوات الاختبار:

أولاً:- المعطيات: يمكن حساب المتوسط الحسابي والانحراف المعياري للمتلاحقات كما يلي:

$$\overline{X}r = \left(\frac{2MN}{M+N}\right)+1$$

$$Sr = \sqrt{\frac{2MN(2MN-M-N)}{(M+N)^2(M+N-1)}}$$

ثانياً:- الفرضيات: فرضية العدم تنص على أنه ليس هناك وجود لاتجاه أو دورات والتوزيع عشوائي فقط لذلك فقيمة Z مساوية للصفر.

H0:Z=0

HA:Z#0

ثالثاً:- الاختبار:

$$Z = \frac{R-\overline{X}r}{Sr}.$$

رابعاً:- تحديد الإقليم الحرج الذي نرفض فيه H0 وإقليم قبولها لإجراء المقارنة ومن ثم الاستنتاج: كما تعودنا نرسم المنحنى ألاعتدالي ونحدد منطقة إقليم رفض H0 بناء على قيمة Z_α المقدرة من الجدول.

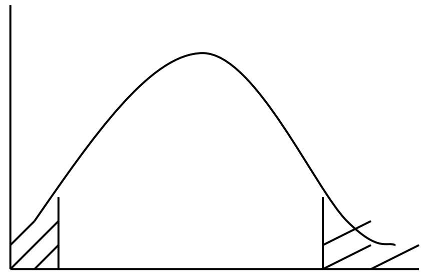
ومن المعطيات المتوفرة نحسب المتوسط الحسابي والانحراف المعياري ثم نجري الاختبار والمقارنة والاستنتاج كما يلي:

$$\overline{X}r = \left(\frac{2x15x10}{15+10}\right)+1=13$$

$$Sr = \sqrt{\frac{2x15x10(2x15x10-15-10)}{(15+10)^2(15+10-1)}} = 2.345$$

$$Z = \frac{11-13}{2.345} = -0.85$$

وبما أن الفرضية البديلة هي لاختبار الطرفين فإن قيمة $Z_{0.05}=$ 1.96 أو 1.96-



وبما إن Z واقعة بإقليم الأمان نقبل H0 ونستنج أن المشاهدات موزعة عشوائيا وليس هناك ما يؤيد وجود أي حياد عن ذلك.

مثال2- قام أحد العلماء بمتابعة حالة زوجته النفسية لمدة 25 يوما ويسجل الرمز g إذا كانت جيدة والرمز b إذا كانت سيئة، والمتلاحقات التالية تبين ذلك:

ggg/bb/g/b/ggg/b/g/bb/gg/b/ggg/bbb/g/b

استخدم α 0.05 لمعرفة هل توزيع حالة الزوجة عشوائية أو تتبع نظام معين يمكن تفسيره.

الحل:

أولاً:- المعطيات والفرضيات:

R=14

Σg=14

Σb=11

H0:Z=0

HA:Z#0

ثانياً:- الاختبار:

$$\overline{X}r=\left(\frac{2x14x11}{14+11}\right)+1=13.32$$

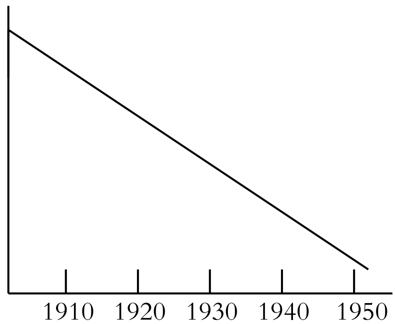
$$Sr=\sqrt{\frac{2x14x11(2x14x11-14-11)}{(14+11)^2(14+11-1)}}=2.41$$

$$Z=\frac{14-13.32}{2.41}=0.282$$

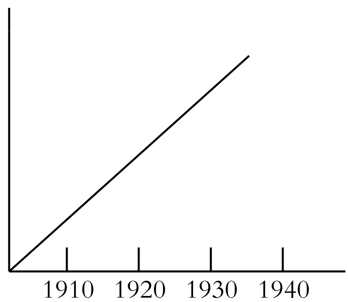
وبما أن **HA** هي لاختبار الطرفين فإن قيمة $Z_{0.05}$=1.96 أو-1.96 وحيث أن Z ليست بإقليم الرفض ،نقبل H0 ونستنتج أن حالة الزوجة ظاهرة عشوائية ولا تتبع أي نظام يمكن تفسيره.

ويستخدم اختبار المتلاحقات كذلك في حالة البيانات الكمية وذلك بتحويل البيانات الأصلية إلى رموز فوق الوسيط أو تحت الوسيط، مثلا بكتابة الرمز a للقيم الأكبر من الوسيط والرمز b للقيم الأصغر منه، ويتم إلغاء القيم المساوية للوسيط.

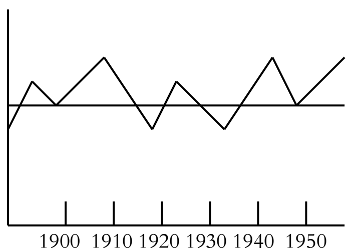
إن أسلوب المتلاحقات يفيد على وجه الخصوص في حالتين رئيسيتين هما اختبار الاتجاهات والأنماط الدورية في المسلسلات الزمنية. فإذا بدأت المتلاحقة برموز أغلبها a (اختصارا لكلمة above أي أكبر من) وانتهت برموز أغلبها b (اختصارا لكلمة below أي أصغر من) فإن هناك اتجاها هابطا، وإذا بدأت برموز أغلبها b وانتهت برموز أغلبها a فإن هناك اتجاها صاعدا. أما إذا كان الرمزان a و b يتبادلان بشكل منتظم فإن هذا يشير إلى وجود نمط عشوائي.



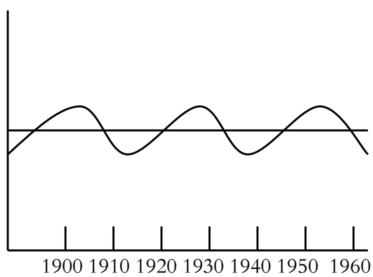
اتجـــاه هابط



اتجـــاه صاعد



نمـط عشـوائي



نمـــط دوري

مثال– الجدول التالي يبين كميات المطر السنوي (ملم) المسجلة خلال 24 سنة، استخدم α 0.05 لاختبار هل المشاهدات عشوائية التوزيع أم تتبع نظاما يمكن تفسيره.

590	540	380	620	460	570	370	430	410	560	290	280
550	520	460	620	350	460	400	570	430	620	360	600

الحل: أولا:– لأجل حساب الوسيط نرتب المشاهدات تصاعديا:

460	460	430	430	410	400	380	370	360	350	290	280
620	620	620	600	590	570	570	560	550	540	520	460

الوسيط = المشاهدة التي تقسم العينة إلى نصفين متساويين وهي460

ثانياً:– نرجع إلى البيانات الأصلية لتحديد موقع كل مفردة من الوسيط، فإذا كانت أكبر منه تأخذ الرمز a وإذا كانت أصغر منه تأخذ الرمز b وتلغى القيم التي تتساوى مع الوسيط. وحيث أن الوسيط =460 نلغي القيم الثلاثة التي تتساوى مع ذلك الرقم من مسلسلة المشاهدات، لكي نحصل على الترتيب التالي للرموز بدلا من القيم الأصلية:

bb/a/bbb/aa/b/aaa/b/a/b/a/bb/aaa

ثالثاً:– المعطيات والفرضيات:

M=10

N=11

R=12

$$Sr = \sqrt{\frac{2x10x11[2x10x11-10-11]}{(10+11)^2(10+11-1)}} = 2.23$$

$$\overline{X}r = \left(\frac{2x10x11}{10+11}\right) + 1 = 11.48$$

رابعاً:– الاختبار:

$$Z = \frac{12-11.48}{2.23} = 0.23$$

خامساً:- المقارنة والاستنتاج:

بما أن HA هي لاختبار الطرفين فإن قيمة $Z_{0.05}$=1.96 و $-$ 1.96 وحيث أن Z الاختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأن توزيع المشاهدات عشوائيا ولا يتبع أي اتجاه أو نسق دوري.

مثال 2- البيانات التالية لنسب الغطاء النباتي لقطاع من الأرض أصابه التملح. أوجد هل توزيع الغطاء النباتي عشوائي أم يظهر نمطا تجميعيا يدل على إصابة الأرض بالتملح. استخدم α 0.05

26	35	30	18	21	28	33	27	27	24
26	28	28	31	19	23	32	29	24	27
21	17	28	23	22	23	22	14	19	24
17	10	21	21	18	20	28	24	13	23

الحل: أولاً:- بعد ترتيب المفردات تصاعديا تبين أن قيمة الوسيط تساوي 23.5

ثانياً:- نرجع للبيانات الأصلية ونقارن قيمها بالوسيط وتستبدل كل قيمة بالرمز a إذا كانت أكبر من 23.5 أو الرمز b إذا كانت أصغر من 23.5

aaaaa/bb/aaaaaaa/bb/aaaaa/bbbbbb/a/bbbb/aa/bbbbbb

من الملاحظ أن القيم الأكبر من الوسيط تميل إلى التجمع في النصف الأعلى من المتلاحقة وأن القيم الأصغر منه تميل إلى التجمع في النصف الأخير. غير أن الحكم الموضوعي على ذلك سيحدد على أساس الاختبار.

ثالثاً:- المعطيات والفرضيات:

M=20

N=20

R1=10

H0:Z=0

HA:Z#0

H0 تنص على أن النمط عشوائي، أما HA فهي تنص على أن النمط تجميعي يدل على التملح.

رابعاً:- تقدير قيمتي المتوسط الحسابي والانحراف المعياري لاستخدامهما في حساب الاختبار Z:

$$\overline{X}r = \left[\frac{2x20x20}{20+20}\right] + 1 = 21$$

$$Sr = \sqrt{\frac{2x20x20(2x20x20-20-20)}{(20+20)^2(20+20-1)}} = 3.12$$

$$Z = \frac{10-21}{3.12} = -3.365$$

خامساً:- المقارنة والاستنتاج:

بما أن HA هي لاختبار الطرفين فإن قيمة $Z_{0.05}$=1.96 أو − 1.96 وحيث أن Z الاختبارية واقعة بالإقليم الحرج نرفض H0 ونستنتج بأن هناك نمطا تجميعيا للغطاء النباتي يدل على استفحال مشكلة التملح.

تمرينات:

1- فيما يلي متلاحقات مسلسلة الجريان السنوي لنهر التيمز عند مدينة تيدينجتون خلال الفترة 1915-1964 ، الرمز a يمثل قيم الجريان الأكبر من الوسيط أما الرمز b فيمثل القيم الأكبر منه. أوجد عند مستوي دلالة 0.95 هل هناك اتجاه في مسلسلة الجريان السنوي لنهر التيمز.

aaa/b/a/bbb/aaaaaa/b/aaa/bbb/aa/b/aaa/bbbb/aa/bbb/a/bb/a/bbb/
a/b/aa/bbb

2- الجدول التالي يبين الأعداد السنوية للطيور المهاجرة التي شوهدت عند أحد المستنقعات خلال 25 سنة متتالية، عند مستوى ثقة 0.95 أختبر هل أعداد الطيور عشوائية أم تظهر اتجاها يمكن تفسيره.

440	370	488	362	510	248	410	380	310	540	395	480	360
-	378	620	405	360	440	524	320	489	330	510	380	398

3- فيما يلي جدول يبين المتوسط المساحي لأعداد العواصف الغبارية بالمنطقة الشمالية الغربية من ليبيا للفترة 1965-1997. أوجد هل توزيع المشاهدات عشوائيا أم يظهر اتجاها نحو الزيادة، 0.05α

16.8	18.7	18.2	16.5	15.4	18.2	20.8	21.9	22.2	24.3	23.0
17.8	18.3	16.5	18.3	19.7	21.4	17.5	16.7	16.9	24.8	19.5
25.9	25.8	24.8	23.6	25.8	24.5	24.2	22.0	21.9	20.8	19.8

4- كثيرا ما تحصل تغيرات في بيانات الرصد الجوي لأحد سببين إما نتيجة لتغيرات طبيعية في المناخ أو نتيجة لتغير موضع محطة الرصد نفسها أو نتيجة لتغير الظروف المكانية بمحيط المحطة. فقد يغير مكان المحطة بسبب وقوعها في مسار طريق من المزمع إنشاؤه أو بسبب التوسع في المطارات أو المنشئات والمباني القريبة منها أو الواقعة في مجال توسعها. وفي أحيان أخرى لا تنقل محطة الرصد من مكانها لكن البيئة المحيطة بها تتعرض لتغيرات ناتجة عن نمو أشجار أو بناء مساكن أو منشآت صناعية تؤثر في اتجاه الرياح وكمية الأمطار الساقطة، إلى غير ذلك. لذلك يجب اختبار عشوائية التوزيع قبل إجراء الاختبارات المعلمية. وبناء على ذلك يتم قبول المسلسلات العشوائية وإقصاء المسلسلات غير العشوائية نظرا لأن إدخالها في التحليل سوف يؤدى إلى التوصل إلى استنتاجات خاطئة، كالتأكيد على وجود اتجاه في الوقت الذي لا يوجد فيه ذلك الاتجاه.

الجدول التالي لمسلسلة الأمطار (ملم) بمدينة طرابلس للفترة 1940-1989 ، استخدم α 0.05 لاختبار عشوائية المسلسلة من عدمه.

274	404	508	142	371	441	360	295	265	336	142
282	473	219	305	337	358	264	402	381	259	215
356	347	228	295	150	209	372	179	214	356	217
441	430	423	405	404	402	386	381	358	351	405
					582	560	552	520	507	473

3- اختبار حسن المطابقة(Goodness of fit-test)

يبحث هذا النوع من التحليل فيما إذا كانت العينة متطابقة مع التوزيع الاحتمالي، أو إلى أي مدى مطابقة النسب المشاهدة مع النسب المتوقعة لها. والاختبارات من هذا النوع جميعها من فئة اختبارات الطرف الأيمن(أكبر من).

افرض أنه أمامنا مكعب للنرد بستة أوجه بأرقام من 1 إلى 6 على التوالي، وافرض كذلك أن P1 هي احتمالية الحصول على الوجه الأول، و P2 هي احتمالية الوجه الثاني و P3 هي احتمالية الوجه الثالث...الخ ، عليه فإن فرضية العدم H0 يمكن أن تصاغ كما يلي:

$$H0:P1=1/6,\ P2=1/6,\ P3=1/6,\ P4=1/6,\ P5=1/6,\ P6=1/6$$

مثال- افرض أننا قلبنا قطعة النرد 120 مرة وتحصلنا على التكرارات المبينة في الجدول التالي استخدم α 0.05 لاختبار H0 القائلة بأن التكرارات المشاهدة لا تختلف عن المتوقعة مما يعني أن قطعة النرد ليست مشحونة إلى جانب دون آخر، أي متوازنة.

جدول (50) التكرارات المشاهدة من رمي قطعة النرد 120 مرة.

الوجه	1	2	3	4	5	6
التكرار	17	25	17	23	15	23

ومن المتوقع أن يحصل كل وجه من الوجوه على(120×1\6= 20 مرة)، وأي اختلاف عن ذلك يرجع إلى أخطاء التعيين. إن الحيود الكبير عن التكرار المتوقع الذي هو (20 مرة) سوف يجعلنا نشتبه في نزاهة قطعة النرد، فقد تكون محملة على هذا الجانب أو ذلك مما يجعلها تستقر عليه أكثر من غيره، وعليه لا يمكنا إرجاع الاختلاف الكبير عن التوزيع المتوقع إلى الصدفة فقط. نحن نواجه بسؤالين: الأول، كيف يمكننا قياس هذا الحيود، وثانيا، كم هو حجم الحيود الذي يجعلنا نرفض H0 .

الاختبار الإحصائي:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

حيث أن: χ^2 كآي تربيع.

O_i التكرار المشاهد على الوجه i.

E_i التكرار المتوقع على الوجه i.

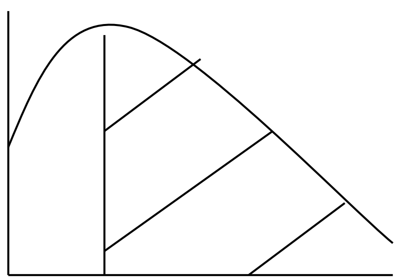
إن اختبار كآي تربيع يقدم لنا مقياس لمدى حيود التكرارات المشاهدة عن المتوقعة (قيمة البسط في المعادلة وهذه القيمة سوف تكون كبيرة إذا كانت القيم المشاهدة (O) متباعدة عن القيم المتوقعة (E)، وسوف تكون صغيرة إذا كانت التكرارات المشاهدة متقاربة مع نظيراتها المتوقعة.

الحل:

أولاً:- نكون جدول لحساب قيمة الاختبار كآي تربيع كما يلي:

الفئة	O	$E_i = NP_i$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
1	17	20	9	0.45
2	25	20	25	1.25
3	17	20	9	0.45
4	23	20	9	0.45
5	15	20	25	1.25
6	23	20	9	0.45
				المجموع=4.30

وهكذا فإن قيمة كآي تربيع =4.30 ، ولتقدير كآي تربيع الجدولية نحتاج إلى درجات الحرية وقيمة α . أما درجات الحرية df فهي تساوي عدد الفئات K مطروح منها 1 أي (df=K-1). والخطوة التالية تتمثل في تقدير قيمة كآي تربيع المعيارية باستخدام درجات الحرية 5 و α لاختبار الطرف الأيمن، ومن الجدول يتبين أن هذه القيمة تساوي 11.07، ألآن نرسم المنحنى ونبين عليه المنطقة الحرجة لأجل المقارنة:



بما أن كآي تربيع χ^2 =4.30 ليست بالإقليم الحرج نقبل H0 ونستنتج بأن التوزيعات المشاهدة لا تختلف اختلافا كبيرا عن التكرارات المتوقعة مما يدل على أن قطعة النرد متوازنة.

مثال 2- يعتقد أن نسب السكان بفصائل الدم A، B، O، و AB هي كما يلي:0.4، 0.2 ، 0.3 و 0.1 على التوالي. وعندما تم اختبار 400 شخص عشوائيا تبين أنهم مقسمين حسب ترتيب الفصائل المذكورة إلى الأعداد التالية:148 ، 96 ، 106 و 50 . استخدم α 0.05 لاختبار H0 التي تؤيد هذا الاعتقاد.

الحل:

أولاً:- المعطيات والفرضيات:

K=4

df=4-1=3

$\alpha = 0.05$

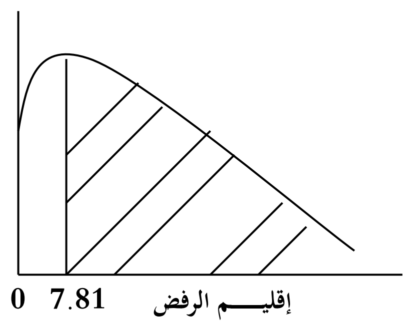
H0:PA=0.4 PB=0.2 PO=0.3 PAB=0.1

HA:PA>0.4 PB>0.2 PO>0.3 PAB>0.1

ثانياً:- حساب قيمة كاي تربيع:

الفئة/فصيلة الدم	O	$E = NP_i$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
A	148	160	144	0.90
B	96	80	256	3.20
O	106	120	196	1.63
AB	50	40	100	2.50
				المجموع =8.23

ثالثاً:- تحديد الإقليم الحرج تحت المنحنى والمقارنة والاستنتاج: باستخدام df =3 و α 0.05 يتبين أن χ^2 = 7.81 ، الآن نرسم المنحنى ونبين عليه الإقليم الحرج لأجل المقارنة:



ومن المقارنة يتبين أن كاي تربيع واقعة بالإقليم الحرج لذلك نرفض H0 ونستنتج أن توزيع فصائل الدم المشاهدة لا يتمشى مع التوزيع المتوقع ربما لوقوع أخطاء في التعيين.

الخلاصة: لإنجاز هذا الاختبار يجب إتباع الخطوات التالية:

1- قسم مجتمع الدراسة إلى عدد K من الفئات هي C1 ، C2 ، C3 ...الخ.

2- فرضية H0 تنص على أن احتمال أن يكون الفرد منتميا إلى الفئة C1 هو P1 ،واحتمال انتمائه للفئة الثانية C2 هو P2 ...الخ.

3- سجل التكرارات المشاهدة في العينة وهي O1 ، O2 ، O3 ...الخ.

4- احسب التكرارات المتوقعة E1 ، E2 ، E3 ...الخ.

5- احسب انحرافات التكرارات المشاهدة عن المتوقعة ثم ربع الناتج $(O-E)^2$.

6- اقسم حاصل الفقرة رقم 5 على E المقابلة لها.

7- اجمع نواتج الخطوة رقم 6 لتحصل على قيمة كآي تربيع.

أمثلة أخرى:

1-لاحظ باحث زراعي بأن مشتله المتكون من 172 شتلة مقسمة كالتالي: 38 شتلة صفراء و134 شتلة خضراء، اختبر الفرضية القائلة بأن نسبة الشتول الصفراء إلى الخضراء هي 1 : 3 .

الحل:

لاحظ أن مجموع النسب 1+3=4

أولاً:- المعطيات و الفرضيات:

K=2

df=1

H0:P1=1/4 , P2=3/4

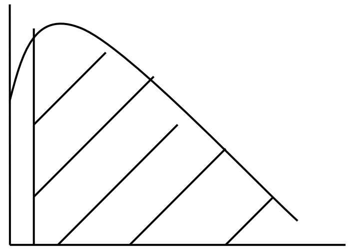
HA:P1>1/4 , P2>3/4

ثانياً:- نكون جدول لحساب قيمة كآي تربيع:

الفئة	O	E=nP	$\frac{(O-E)^2}{E}$
الصفراء	38	43	0.5813
الخضراء	134	129	0.1937
			المجموع=0.77

ثالثاً:- تحديد الإقليم الحرج تحت المنحنى لأجل المقارنة :

قيمة كآي تربيع ألمقدره عند 0.05α ودرجة حرية واحدة = 3.841 ، وبما أن كآي تربيع الإختبارية ليست بالإقليم الحرج نقبل H0 ونستنتج بأنه ليس هناك فرق بين ما شاهده الباحث وبين النسبة الثابتة.



2- من تعداد عام 1960 حسبت احتمالات أعمار فئات البالغين الذين تزيد أعمارهم عن 20 سنة ورتبت في العمود الأول والثاني من الجدول التالي، وفي عام 1976 تمت مقابلة عينة أخري من البالغين ورتبت تكراراتهم في العمود الثالث. اختبر الفرضية القائلة بأن أعمار 1976 لا تختلف عن أعمار 1960، استخدم 0.05α

جدول (51) فئات السن واحتمالات أعمار تعداد 1960 وتكرار أعمار 1976

فئات السن C لعام 1960	P_{1960}	O_{1976}
20-29	0.25	270
30-39	0.23	202
40-49	0.20	180
50-59	0.15	160
60-69	0.08	88
> 70	0.09	100
		مج=1000

الحل:

أولاً:- المعطيات والفرضيات:

K=6

N=1000

df=6-1=5

$\alpha = 0.05$

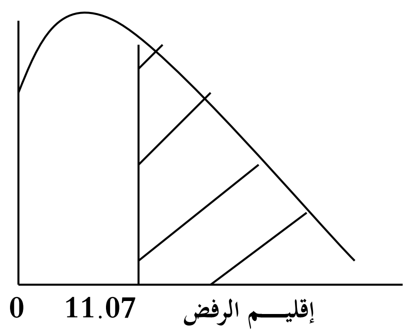
H0:P1=0.25 , P2=0.23 , P3=0.20 , P4=0.15 , P5=0.08 , P6=0.09

HA:P1>0.25 ,P2>0.23 , P3>0.20 , P4>0.15 , P5>0.08 , P6>0.09

ثانياً:- حساب قيمة الاختبار كآي تربيع:

Cالفئات	O	E=nP	$\frac{(O-E)^2}{E}$
C1	270	250	1.6
C2	202	230	3.4
C3	180	200	2.0
C4	160	150	0.66
C5	88	80	0.80
C6	100	90	1.11
			المجموع=9.57

ثالثاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج: قيمة كآي تربيع المقدرة عند مستوى دلالة 0.95 و5 درجات حرية = 11.07 ومن المقارنة مع القيمة الاختبارية (أنظر الشكل) يتبين أن الاختبارية ليست بالإقليم الحرج لذلك نقبل H0 ونستنتج بأن تكرارات أعمار 1976 لا تختلف اختلافا جذريا عن توزيع أعمار 1960.



3- آلة يفترض أنها تخلط ثلاثة أنواع من الحلوى (الكرميل والتوفي و الشيكولاتة) بالنسب 4: 3: 2. ومن دراسة عينة من 270 قطعة حلوى تبين أنها تتكون من 135 كرميل و 70 توفي و 65 شيكولاتة. المطلوب اختبار H0 القائلة بأن الآلة تخلط الحلوى بالنسب المذكورة أعلاه. α 0.05

الحل:

أولاً:- المعطيات والفرضيات:

K=3

N=270

$\Sigma P=4+3+2=9$

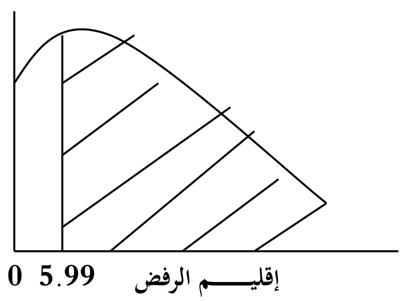
$H0:P1=4/9\ ,\ P2=3/9\ ,\ P3=2/9$

$HA:P1>4/9\ ,P2>3/9\ ,\ P3>2/9$

ثانياً:- حساب قيمة كآي تربيع:

الفئة C	O	E=nP	$\frac{(O-E)^2}{E}$
C1	135	120	1.875
C2	70	90	4.444
C3	65	60	0.416
			المجموع=**6.74**

ثالثاً:- تحديد الإقليم الحرج والمقارنة والاستنتاج: قيمة كآي تربيع المقدرة من الجدول المعياري عند مستوى 0.95 ودرجتان للحرية= 5.99 ، وحيث أن كآي تربيع الاختبارية أكبر من الجدولية أي واقعة بالإقليم الحرج نرفض H0 ونستنتج بأن الآلة لا تخلط الحلوى بالنسب المذكورة.



تمرينات:

1- فيما يلي أعداد الزبائن وأعداد الأيام التي ترددوا خلالها على المحاسب لعينة من 400 يوم عشوائي:

عدد الزبائن	0	1	2	3	4	5	6
عدد الأيام	50	112	125	60	28	16	9

أوجد هل هناك حيود عن التوزيع النظري المبين في الجدول التالي. استخدم 0.05α

عدد الزبائن	0	1	2	3	4	5	6
عدد الأيام	0.10	0.27	0.27	0.18	0.09	0.05	0.04

2- لقد تمت مشاهدة 500 سيارة على الطريق السريع منها 160 بيجو و 100 تويوتا و 80 مازدا و 25 مرشيدس و 130 فولفو. هل التكرارات المشاهدة متمشية مع التكرارات المتوقعة بالنسب التالية 30 : 15 : 10 : 5 : 20 .استخدم 0.05α

4- اختبار الارتباط للبيانات الوصفية: Spearman-RankCorrelation

اشتق الباحث سبيرمان هذا المعامل (r_s) لاستخدامه في حالة البيانات الوصفية، كما يمكن استعماله أيضا في حالة البيانات الكمية، إلا أن نتائجه أضعف من معامل ارتباط بيرسون (r). إن أحد أهم فرضيات تحليل الارتباط بين متغيرين تتمثل في افتراض التوزيع الاعتدالي للمشاهدات، وفي حالة مخالفة هذه الفرضية لا يجدي استعمال ارتباط بيرسون لكن من الممكن استعمال ارتباط سبيرمان للرتب الذي لا تتأثر نتائجه بهذه المخالفة.

يعتمد ارتباط الرتب على رتبة المفردة بدلا من قيمتها وذلك بعد ترتيب المشاهدات تنازليا أو تصاعديا. ففي كثير من الأحيان نجد بيانات عن تقديرات الامتحانات أو تقديرات الإنتاج الزراعي أو شدة المطر أو سرعة الرياح أو ارتفاع الأمواج أو معدل التعرية...الخ. ففي هذه الأحوال وغيرها نستبدل الوصف بالرتبة التي تحتل بدورها محل المشاهدات الأصلية. إذ ينبغي علينا في البداية إعطاء رتب للمشاهدات النوعية على النحو الذي يتناسب مع درجة المشاهدة وأهميتها. فعلى سبيل المثال، في تقديرات الامتحانات نجد ممتاز جدا، ممتاز، جيد جدا، جيد، مقبول، ضعيف، وضعيف جدا. فإذا كان الترتيب تنازليا أصبحت الرتب:1، 2، 3، 4، 5، 6، و 7 على اعتبار أن رتبة ممتاز جدا هي 1 أما رتبة ضعيف جدا فهي 7. وبالنسبة لسرعة الرياح نجد أن هناك رياح هوجاء، عاصفة، وسريعة، ومتوسطة وضعيفة وساكنة وتكون رتبها 1 للهوجاء و 6 للساكنة.

ويحسب معامل ارتباط سبيرمان للرتب بالمعادلة التالية:

$$r_s = 1 - \left[\frac{6\left(\sum d^2\right)}{n\left(n^2 - 1\right)} \right].$$

حيث أن: d^2 = مربع فرق الرتب المتناظرة للمتغيرين x و y .

n = عدد المفردات المتناظرة للمتغيرين x و y .

وعندما يكون حجم العينة n أكبر من 10 يمكن اختبار الدلالة الإحصائية لمعامل الارتباط باستخدام الاختبار التائي كما يلي:

$$T = r_s \sqrt{\frac{n-2}{1-r_s^2}}$$

وتقارن قيمة T بقيمة T_α المقدرة بدرجات الحرية (df=n-2).

أمثلة:

1- الجدول التالي يبين تقديرات 6 طلاب في مادتي المناخ والخرائط، هل هناك ارتباط قوي بين درجات الطالب في المادتين.

رقم الطالب	1	2	3	4	5	6
درجة المناخ	ممتاز	مقبول	جيد جدا	ضعيف	ضعيف جدا	جيد
درجة الخرائط	جيد	مقبول	ممتاز	ضعيف جدا	جيد جدا	ضعيف

الحل:

أ- نرتب التقديرات تنازليا لمعرفة رتبة كل مفردة كما يلي:

التقديرات	ممتاز	جيد جدا	جيد	مقبول	ضعيف	ضعيف جدا
الرتبة	1	2	3	4	5	6

ب- نكون جدول لحساب عناصر معادلة الارتباط كما يلي:

d^2	d	رتب y	رتبx	الخرائطy	المناخx
4	-2	3	1	جيد	ممتاز
0	0	4	4	مقبول	مقبول
1	1	1	2	ممتاز	جيد جدا
1	-1	6	5	ضعيف جدا	ضعيف
16	4	2	6	جيد جدا	ضعيف جدا
4	-2	5	3	ضعيف	جيد
المجموع=26					

ج- نحسب قيمة الارتباط بالمعادلة كما يلي:

$$r_s = 1 - \left[\frac{6(26)}{6(6^2 - 1)} \right] = 0.25$$

د- الاستنتاج هناك ارتباط موجب لكنه ضعيف بين درجات الامتحانين.

2- افرض أن هناك حكمين على درجة أداء 7 لاعبي جمباز، ومطلوب من كل حكم أن يرتب اللاعبين من 1 إلى 7 وكانت نتائجهم كما هي مبينة في الجدول التالي، احسب معامل ارتباط سبيرمان للرتب لمعرفة هل هناك تناسق بين تقييم الحكمين أم لا.

اللاعب	الحكم الأولx	الحكم الثانيy
1	2	7
2	7	6
3	1	1
4	3	3
5	5	5
6	6	2
7	4	4

الحل:

أولاً:- لاحظ أن الرتب جاهزة وليست كما في المثال السابق، لذلك ننتقل مباشرة إلى حساب $\sum d^2$ وذلك بتكوين جدول كما يلي:

ثانياً:- نحسب معامل الارتباط:

رتبx	رتبy	d	d^2
2	7	-5	25
7	6	1	1
1	1	0	0
3	3	0	0
5	5	0	0
6	2	4	16
4	4	0	0
			المجموع=42

$$r_s = 1 - \left[\frac{6(42)}{7(7^2 - 1)} \right] = 0.25$$

ثالثاً:- الاستنتاج: معامل الارتباط بقيمة0.25 تشير إلى وجود علاقة موجبة لكنها ضعيفة بين تقديرات الحكمين مما يدل على أنهما ليسا متفقان تماما في تقديراتهما للاعبين.

3- تقدم 10 أفراد لشغل وظيفة أستاذ بالجامعة وثم ترتيبهم حسب درجة امتحان المقابلة مع كل من العميد ورئيس القسم كما يلي:

رقم المتقدم	1	2	3	4	5	6	7	8	9	10
رتب رئيس القسم	9	5	10	6	8	1	3	7	4	2
رتب العميد	4	3	6	7	10	5	1	8	9	2

المطلوب:

أ- حساب معامل ارتباط الرتب.

ب- اختبار H0 القائلة بعدم وجود علاقة بين رتب رئيس القسم ورتب العميد، استخدم مستوى دلالة 0.95

الحل: أولاً:- الفرضيات.

$$H0: r_s = 0$$
$$HA: r_s \# 0$$

HA تقول بوجود علاقة ارتباط قوية بين رتب رئيس القسم والعميد.

ثانياً:- لحساب معامل الارتباط نكون جدولا:

d^2	d	رتب العميدy	رتب رئيس القسمx
25	5	4	9
4	2	3	5
16	4	6	10
1	-1	7	6
4	-2	10	8
16	-4	5	1
4	2	1	3
1	-1	8	7
25	-5	9	4
0	0	2	2
المجموع=96			

$$r_s = 1 - \left[\frac{6(96)}{10(10^2 - 1)} \right] = 0.42$$

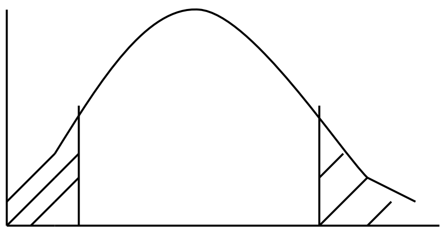
لاحظ أن هناك ارتباط موجب لكنه ضعيف وللتأكد من دلالته الإحصائية نستخدم اختبار T.

رابعاً:- الاختبار الإحصائي:

$$T = 0.42\sqrt{\frac{10-2}{1-0.42^2}} = 1.3$$

خامساً:- المقارنة والاستنتاج:

قيمة $T_{0.05}$ لاختبار الطرفين = 2.3 أو -2.3 وبما أن T(1.3) ليست بالإقليم الحرج نقبل H0 ونستنتج بأنه ليس هناك علاقة ذات دلالة إحصائية بين رتب رئيس القسم ورتب العميد.



تمرينات:

1- فيما يلي بيانات افتراضية لكميات المطر (ملم) وكميات القمح (قنطار/هكتار) لإحدى عشر سنة متتالية، احسب r_s واختبر دلالته عند مستوى 0.95

كمية المطر	100	200	350	150	210	450	380	500	160	470	140
كمية القمح	10	30	35	15	21	45	38	50	16	47	14

2- الجدول التالي لبيانات افتراضية لمعدل الارتفاع ومعدل درجة الحرارة، والمطلوب حساب معامل الارتباط واختبار دلالته الإحصائية عند مستوى 0.95

0	200	400	300	500	800	450	1000	1500
30	25	20	19	17	14	9	7	5

3- الجدول التالي يبين سرعة الرياح بالعقدة وارتفاع الموج بالأمتار المسجلة بأحد الموانئ، احسب r_s، واختبر دلالته الإحصائية عند معدل 0.95

سرعة الرياح بالعقدة	10	50	15	30	20	60	100	5	120	35	40	80
ارتفاع الموج بالأمتار	1	4.5	1.5	3	2	5	9	0.3	15	3	3.5	7

4- الجدول التالي يبين حالات الملاريا السنوية وكميات الأمطار المسجلة بزمبابوي. اختبر هل هناك علاقة ارتباط بين تفشى الملاريا وكميات الأمطار(ملم)،مستوى الدلالة 0.95 .

السنة	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
المطر	450	750	600	620	500	330	630	500	400	700	800
الملاريا	42000	96000	70000	60000	58000	40000	81000	78000	72000	17000	188000

الفصل السادس

تحليل التباين

مقدمة:

لقد سبق وأن بيننا الاختبار التائي لمقارنة متوسطي مجتمعين في حالة عدم معرفة تباينهما على افتراض أن المتوسطين (u1 و u2) متساويين. والآن نعمم الفكرة لأكثر من مجتمعين، فعلى سبيل المثال، قد نرغب في مقارنة عدد من الأسمدة الكيماوية، أو عدة أساليب صناعية أو طرق تدريس أو تقنيات إعلانية...الخ. وكالمعتاد سوف نختار عينات من هذه المجتمعات، وفي الغالب نتوقع أن تكون المتوسطات صغيرة لدرجة يمكن إرجاعها إلى الصدفة فقط أم أنها اختلافات كبيرة ولا يمكن إرجاعها إلى الصدفة، بل هي ناتجة عن حقيقة أن العينات مأخوذة من مجتمعات مختلفة وليست متجانسة.

الأسلوب المتبع في هذه الحالة هو تحليل التباين ويختصر في (ANOVA) الذي يستخدم اختبار واحد لمقارنة العديد من المتوسطات في آن واحد. ويتطلب هذا النوع من التحليل الاستقرائي الإيفاء بأربعة شروط أو فرضيات هي:

1-العشوائية في اختيار مفردات العينة، أي تساوي فرص أو احتمالات اختيار المفردات، وبمعنى آخر ليس هناك انحياز في اختيار المفردات.

2- اختيار العينة من مجتمع اعتدالي.

3- تساوي التباين، أي ليس هناك اختلاف في حجم التباين بين فئات العينة أو مستوياتها (أي لا يظهر زيادة في فترة وتناقص في فترة أخرى نتيجة لوجود دورات أو اتجاه).

4- استقلال العينات المكونة لعينة الدراسة، أي لا يوجد ارتباط بينها، كما أن مفردات كل عينة ليس لها علاقة ارتباط ببعضها الآخر كالحال عند وجود اتجاه نحو الزيادة أو التناقص.

إن استيفاء هذه الشروط ضروري لنجاح هذا التحليل، وللتحقق من ذلك لا بد من تعريض البيانات للتحاليل التي سبق الإشارة إليها في الفصل السابق(ألاعتدالية، الاستقلالية واختبار المتلاحقات)...الخ.

من أين تأتي الأخطاء (errors) وتختصر (e_s): تأتي الأخطاء في العينة من عدة مصادر مثل:

1- اختلاف الأشخاص أو الأشياء الداخلة في التجربة حتى عندما تعرض لمؤثرات ثابتة. إذ يحصل ذلك نتيجة اختلاف الظروف الثقافية واختلاف مستوى الذكاء بين الأفراد، فالأشخاص الذين يدرسون ساعات متساوية من الزمن سوف لن يحصلوا على نفس الدرجة في الامتحان نتيجة لوجود اختلاف في مستويات الذكاء لديهم ولاختلاف أنماط عيشهم وخلفياتهم العلمية والثقافية والاجتماعية.

2- أخطاء في القياس: قد تحصل أخطاء أثناء محاولتنا قياس أو امتحان الطلبة، ربما بسبب عدم دقة الأجهزة. فقد تكون الطباعة مشوهة في هذه الورقة أو تلك أو أن الرقم الصحيح غير موجود فيقرأه هذا الطالب بطريقة تختلف عن غيره. مثلا، تأمل الفقرة التالية: "للحياة في المدينة إيجابيات وسلبيات" وضح ذلك؟. أحد الطلبة قرأها كما يلي (للمياه في المدينة إيجابيات وسلبيات) فكانت إجابته بعيدة كل البعد عن الهدف المطلوب. وكثيرا ما يساء فهم الأرقام، فمثلا بدلا من 50 مليمتر (ملم) يقرأها البعض 50 سنتمتر. وقد تحصل أخطاء نتيجة لعدم ملائمة النموذج الإحصائي، فقد يكون هذا النموذج معد لاختبار علاقة خطية، لكن العلاقة في البيانات أسية أو لوغاريثمية. وقد يحصل الخطأ نتيجة لوضع متغير غير مناسب في المعادلة، أو نتيجة لإهمال متغير ضروري وعدم إدراجه فيها.

3- أخطاء طباعة، فبدلا من طباعة الرقم 27 يطبع 72 والفرق بينهما شاسع وبدلا من طباعة 2.7 يطبع 7.2 مما يعطي استنتاجات خاطئة خاصة في مواضيع حساسة كالتحاليل الطبية مثلا فيوصف الطبيب دواء مكلف وغير ضروري أصلا.

يوجد أسلوبين لتحليل التباين هما التحليل الأحادي والتحليل الثنائي للتباين نكتفي هنا بشرح الأول منهما.

تحليل التباين الأحادي:

أمثلة:

1- تريد شركة لإيجار السيارات مقارنة ثلاثة أنواع من البنزينA، B و C، افرض أنها أجرت أربعة تجارب على كل نوع بحيث تقيس عدد الكيلومترات التي يسيرها الجالون من البنزين في كل مرة وبوبت النتائج في الجدول التالي:

جدول(52) عدد الكيلومترات التي يسيرها الجالون من البنزين

نـــــوع البنـــزين		
A	**B**	**C**
15	19	22
19	17	17
14	16	19
16	20	18
$\overline{Xa}=16$	$\overline{Xb}=18$	$\overline{Xc}=19$

من الجدول يتبين أن القراءات الفردية متباينة وكذلك متوسطات الأنواع الثلاثة. فهل هذه الاختلافات في متوسطات الأنواع راجعة إلى عامل الصدفة أم أنها راجعة إلى حقيقة اختلاف الأنواع من البنزين. وبعبارة أخرى إذا كانت u1 و u2 و u3 تمثل متوسطات الكيلومترات التي يسيرها الجالون في العينة 1 و 2 و 3 على التوالي فإن فرضية العدم التي نرغب في اختبارها تقول بأن متوسطات العينات الثلاثة متساوية أما الفرضية البديلة فهي تقول أن المتوسطات غير متساوية:

H0:u1=u2=u3

HA:u1#u2#u3

واضح أنه إذا كان اختلاف متوسطات العينات صغير جدا فسوف نقبلH0، لكن إذا كان هذا الاختلاف كبيرا فسوف نرفض H0 . فإذا كان الاختلاف صغيرا يمكن إرجاع ذلك لعامل الصدفة ونقرر بأن متوسطات الأنواع التي هي 16- للنوع الأول و 18- للنوع الثاني و 19- للنوع الثالث ما هي إلا ثلاث نقاط تقديرية للمتوسط العام $\overline{u}$.

إن تحليل التباين لا يقتصر على مقارنة متوسطات العينات فقط بل يهتم أيضا بتحليل التباين داخل العينات نفسها. لاحظ من الجدول التالي أن مدى التباين الداخلي كبير بالرغم من تقارب متوسطات العينات.

جدول (53) التباين داخل العينات

A	B	C
19	14	13
15	19	15
24	21	19
18	18	17
u=19	u=18	u=16

وعند استخدامنا للمعادلات الإحصائية نستعين بإشارات تتبع المفردات لكي تحدد مواقعها بالنسبة للعينة وبالنسبة للمجتمع ككل. فالرقم الواقع إلى يمين X مباشرة يوضح رقم العينة، أما الرقم الثاني فيبين موقع المفردة في العينة نفسها. مثلاً ($X_{1,2}$) تعني المفردة الثانية من العينة الأولى، أما ($X_{2,4}$) فتعني المفردة الرابعة من العينة الثانية...الخ. لاحظ الجدول التالي:

جدول(54) إشارات تتبع المفردات داخل كل عينة

A	B	C
$X_{1,1}$ $X_{1,2}$ $X_{1,3}$ $X_{1,4}$	$X_{2,1}$ $X_{2,2}$ $X_{2,3}$ $X_{2,4}$	$X_{3,1}$ $X_{3,2}$ $X_{3,3}$ $X_{3,4}$
$\overline{X}_A = \frac{\Sigma X_1}{n}$	$\overline{X}_B = \frac{\Sigma X_2}{n}$	$\overline{X}_c = \frac{\Sigma X_3}{n}$

ويحسب **<u>المتوسط العام للمتوسطات</u>** بقسمة مجموع المتوسطات على عددها ويرمز له $\overline{\overline{X}}$

$$\overline{\overline{X}} = \frac{\overline{X}_A + \overline{X}_B + \overline{X}_C}{n} \quad \text{.........(16)}$$

ولكل فئة أو مجموعة أو نوع يحسب تباينه بأي من المعادلتين التاليتين:

$$S_1^2 = \frac{\Sigma\left(X_1 - \overline{X_1}\right)^2}{n-1}$$

$$S_1^2 = \frac{\Sigma X_1^2 - \frac{\left(\Sigma X_1\right)^2}{n}}{n-1}$$

ويحسب تباين العينة الثانية والثالثة بنفس الطريقة. وجميع هذه التباينات تعتبر تقديرات غير منحازة لتباين المجتمعات المأخوذة منها $\sigma_3^2, \sigma_2^2, \sigma_1^2$. وبما أننا نفترض أن جميع المجتمعات تحتوي على نفس التباين σ^2 فمن الممكن الحصول على تقدير متطور للتباين بترجيح تباينات العينات الثلاثة:

$$S_P^2 = \frac{S_1^2 + S_2^2 + S_3^2}{3}$$

وباستعمال إشارات الجمع المزدوجة فإن **التباين المرجح** يمكن أن يحسب بالصيغة المختصرة التالية:

التباين المرجح: $$S_p^2 = \frac{\sum^{j}\sum^{n}(X_{ij} - \overline{X})^2}{J(n-1)}$$

حيث أن: j = عدد العينات(3)

N = حجم العينة (4)

إذن التباين المرجح S_P^2 هو أفضل تقدير للتباين العام للمجتمع σ^2.

إن بسط المعادلة $\sum^{j}\sum^{n}(X_{ij} - \overline{X_i})$ ما هو إلا المجموع ألتربيعي للأخطاء(البواقي) الذي إذا قسمناه على درجات الحرية(j(n-1 يعطينا تقدير للتباين العام σ^2 ، وبالنسبة لجدول عدد الكيلومترات التي يسيرها جالون البنزين(جدول 2) يمكن أن نحسب المجموع التربيعي للأخطاء كما يلي:

جدول(55) حساب المجموع التربيعي للأخطاء (البواقي)..

$(X_1-\overline{X}_1)^2$	$(X_2-\overline{X}_2)^2$	$(X_3-\overline{X}_3)^2$
1	1	9
9	1	4
4	4	0
0	4	1
Σ=14	Σ=10	Σ=14

إذن المجموع التربيعي للبواقي أو الأخطاء والذي يختصر في SSE يحسب كما يلي:

$$SSE=\sum^{j}\sum^{n}(X-\overline{X})^2 \qquad (17)$$

الآن نعوض البيانات من الجدول(4) في المعادلة رقم(3) لحساب SSE :

SSE=14+10+14=38

أما متوسط المجموع التربيعي للبواقي MSSE فيحسب بقسمة SSE على درجات الحرية:

$$MSSE=\frac{SSE}{J(n-1)} \qquad (18)$$

$$MSSE=\frac{38}{3(4-1)}=4.22$$

والخطوة التالية في التحليل تتمثل في حساب المجموع التربيعي بين المتوسطات SSb. وهنا نفترض أن العينات الثلاثة تكون فيما بينها عينة كبيرة أخذت من مجتمع واحد بمتوسط μ وتباين σ^2. وعليه فإن المتوسطات $\overline{X}_1$ و $\overline{X}_2$ و $\overline{X}_3$ تشكل عينة من ثلاث متوسطات أخذت من ذلك المجتمع ومن ثم فإن التباين المرجح للعينات الثلاثة يحسب كما يلي:

$$S_{\bar{X}}^2 = \frac{\left(\bar{X}_1 - \bar{\bar{X}}\right)^2 + \left(\bar{X}_2 - \bar{\bar{X}}\right)^2 + \left(\bar{X}_3 - \bar{\bar{X}}\right)^2}{3-1}$$

$$S_{\bar{X}}^2 = \frac{\sum^{3}\left(\bar{X} - \bar{\bar{X}}\right)^2}{n-1}$$

ويمكن الحصول على تقدير غير منحاز لتباين المجتمع σ^2 بضرب تباين المتوسطات في حجم العينة مقسوما علي درجات الحرية كما يلي:

$$\sigma^2 = \frac{n\sum\left(\bar{X} - \bar{\bar{X}}\right)^2}{j-1}$$

إن قيمة البسط في المعادلة السابقة يسمى المجموع التربيعي بين المتوسطات SSb ولها درجات حرية تساوي عدد العينات مطروح منها 1 أي (J-1).

$$SSb = n\sum\left(\bar{X} - \bar{\bar{X}}\right)^2$$

وإذا قسمنا المجموع التربيعي بين المتوسطات SSb على درجات الحرية نحصل على **متوسط المجموع التربيعي بين المتوسطات MSSb** وهو أفضل تقدير للتباين σ^2.

وفي مثالنا نجد أن المتوسط العام $\bar{\bar{X}}$ يساوي مجموع المتوسطات مقسوما على 3 كما يلي:

$$\bar{\bar{X}} = \frac{16+18+19}{3} = 17.67$$

أما SSb فهي تحسب كما يلي:

$$SSb = n\left[\left(\bar{X}_1 - \bar{\bar{X}}\right)^2 + \left(\bar{X}_2 - \bar{\bar{X}}\right)^2 + \left(\bar{X}_3 - \bar{\bar{X}}\right)^2\right]$$

$$SSb = 4\left[(16-17.67)^2 + (18-17.67)^2 + (19-17.67)^2\right] = 18.667$$

وتحسب mssb بقسمة SSb على درجات الحرية (j-1):

$$MSSb = \frac{SSb}{J-1} = \frac{18.667}{3-1} = 9.33$$

إذا 9.33 هي تقدير غير منحاز للتباين العام على افتراض أن المتوسطات متساوية:

$$H0: \mu_1 = \mu_2 = \mu_3$$

مما تقدم يتبين أن MSSE تعكس قيمة التباين بين المجموعات، إنه التباين حول متوسط كل مجموعة. أماMSSb فهو يعكس تباين المتوسطات حول المتوسط العام وهو الآخر يساوي σ^2. فإذا كان هناك اختلاف بين اثنين على الأقل من متوسطات المجموعات فإنMSSb سوف تكون أكبر من σ^2.

الخلاصة:

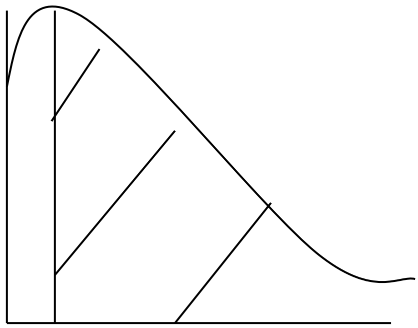
- إذا كانت H0 صحيحة فإن: MSSE = σ^2 و MSSb= σ^2
- أما إذا كانت H0 خاطئة فإن: MSSE= σ^2 لكن MSSb ستكون أكبر من σ^2
- الخطوة الأخيرة في تحليل التباين ANOVA تتمثل في اختبار الفرضيات. فإذا كانت H0 صحيحة فإن التقديرين الاثنين المتحصل عليهما للتباين MSSE و MSSb هما تقديرين صحيحين غير منحازين للتباين σ^2 ويجب أن يكونا متقاربين في القيمة، وللتحقق من ذلك نستخدم اختبارF كما يلي:

$$F = \frac{MSSb}{MSSE}$$

وتقارن قيمة F الاختبارية بقيمة F_α المقدرة من الجدول بدرجات الحرية df التي تساوي (J−1) للبسط و (n−1)J للمقام:

$$F = \frac{9.33}{4.22} = 2.21$$

أما قيمة F_α المقدرة عند $\alpha 0.05$ و درجات حرية = 2 للبسط و 8 للمقام فهي تساوي 4.26 . نرسم المنحنى ونبين عليه موقع $F\alpha$:



وحيت أن F ليست بالإقليم الحرج نقبل H0 ونستنتج بأنه ليس هناك دليل على وجود اختلافات جوهرية في أنواع الوقود الثلاثة.

أمثلة:

1- الجدول التالي لدرجات امتحان أربعة مجموعات من الطلبة، درس كل منها بأسلوب مختلف. كلفت المجموعةA بكتابة إطار عام لنظام النقد الإنجليزي، وكلفت المجموعة B بكتابة ملخص لنظام النقد، وكلفت المجموعة C بدراسة كتاب مبرمج لنظام النقد الإنجليزي أما المجموعة D فكلفت بمشاهدة شريط عن نظام النقد الإنجليزي. اختبر H0 القائلة بعدم وجود اختلافات جوهرية في نتائج الأساليب الأربعة، استخدم معدل ثقة (0.95).

جدول (56) درجات امتحان أربع مجموعات من الطلبة درسوا بأسلوب مختلف.

A	B	C	D
26	51	52	41
34	50	64	49
46	33	39	56
48	28	54	64
42	47	58	72
49	50	53	65
74	48	77	63
61	60	56	87
51	71	63	77
53	42	59	62

الحل:

أولاً:- الفرضيات:

H0:u1=u2=u3=u4

HA:u1>u2>u3>u4

ثانياً:- نكون جدول لحساب عناصرSSE وSSb :

A $(X_1 - 48.4)^2$		**B** $(X_2 - 48)^2$		**C** $(X_3 - 57.5)^2$		**D** $(X_4 - 63.6)^2$	
26	501.76	51	9	52	30.25	41	510.76
34	207.36	50	4	64	42.25	49	212.16
46	5.76	33	225	39	342.25	56	57.76
48	0.16	28	400	54	12.25	64	0.16
42	40.96	47	1	58	0.25	72	70.56
49	0.36	50	4	53	20.25	65	1.96
74	655.36	48	0	77	380.25	63	0.36
61	158.76	60	144	56	2.25	87	547.56
51	6.76	71	529	63	30.25	77	179.56
53	21.16	42	36	59	2.25	62	2.56
ΣA=484		ΣB=480		ΣC=575		ΣD=636	

$$\sum (X_1 - 48.4)^2 = 1598.5$$

$$\sum (X_2 - 48)^2 = 1352$$

$$\sum (X_3 - 57.5)^2 = 762.51$$

$$\sum (X_4 - 63.6)^2 = 1584.4$$

ثالثاً:- حساب المجموع التربيعي للبواقي SSE والمتوسط العام للمجموع التربيعي للبواقي : MSSE

SSE=1598.5+1352+762.51+1584.4=5297.5

$$MSSE = \frac{5297.41}{4(10-1)} = 147.5$$

رابعاً:- حساب المجموع التربيعي بين المتوسطات SSb و المتوسط العام للمجموع التربيعي بين المتوسطات MSSb :

يحسب متوسط المتوسطات بجمع المتوسطات وقسمتها على عدد المجموعات:

$$\overline{\overline{X}} = \frac{48.4 + 48 + 57.5 + 63.6}{4} = 54.35$$

أما المجموع التربيعي لانحرافات المتوسطات عن المتوسط العام فيحسب بالمعادلة التالية:

$$SSb = 10\left[(48.4 - 54.35)^2 + (48 - 54.35)^2 + (57.5 - 54.35)^2 + (63.6 - 54.35)^2\right] = 1724.1$$

ويحسب متوسط المجموع التربيعي لانحرافات المتوسطات عن المتوسط العام (على درجات الحرية.SSb) بقسمة (MSSb)

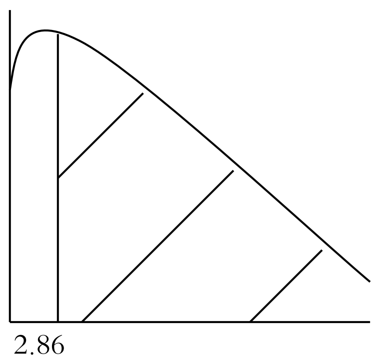
$$MSSb = \frac{1724.1}{(4-1)} = 574.7$$

خامساً:- حساب قيمة الاختبار الإحصائي والمقارنة والاستنتاج:

F=MSSb/MSSE

$$F = \frac{574.7}{147.15} = 3.9$$

أما قيمة $F_{0.05}$ المقدرة بدرجات الحرية 3 للبسط و 36 للمقام فتساوي 2.86



وبما أن F واقعة بالإقليم الحرج نرفض H0 ونستنتج أن اختلاف أسلوب التدريس تسبب في اختلاف النتائج بين مجموعتين على اقل تقدير.

الفصل السابع

تحليل المسلسلة الزمنية

المسلسلة الزمنية هي مجموعة من القيم لظاهرة ما في فترات زمنية متعاقبة طبقا لأزمنة حدوثها، فالمسلسلة تحتوي على متغيرين هما الزمن (T) والظاهرة (y) محل الدراسة.

يمكن تمثيل أي مسلسلة زمنية بيانيا بأخذ قيم الزمن (T) على المحور السيني وقيم الظاهرة (Y) على المحور الرأسي وتحدد نقطة لكل سنة أو فترة زمنية وتوصل النقاط فيتم الحصول حلى خط منكسر يسمى المنحنى التاريخي للسلسة. أي أن السلسلة الزمنية يمكن تمثيلها بنقطة تتحرك عبر الزمن تاركة خطا منكسرا يمثل قيم الظاهرة المدروسة.

مكونات المسلسلة الزمنية:

هناك عدد من المؤثرات المشتركة في كل مسلسلة ولكن بدرجات متفاوتة من ظاهرة إلى أخرى طبقا لطبيعة الظاهرة قيد الدراسة. وتوجد أربعة مؤثرات تطرأ على الظاهرة عبر الزمن هي: 1- الاتجاه العام، 2- التغيرات الموسمية، 3- التغيرات الدورية، و 4- التغيرات العشوائية (العرضية).

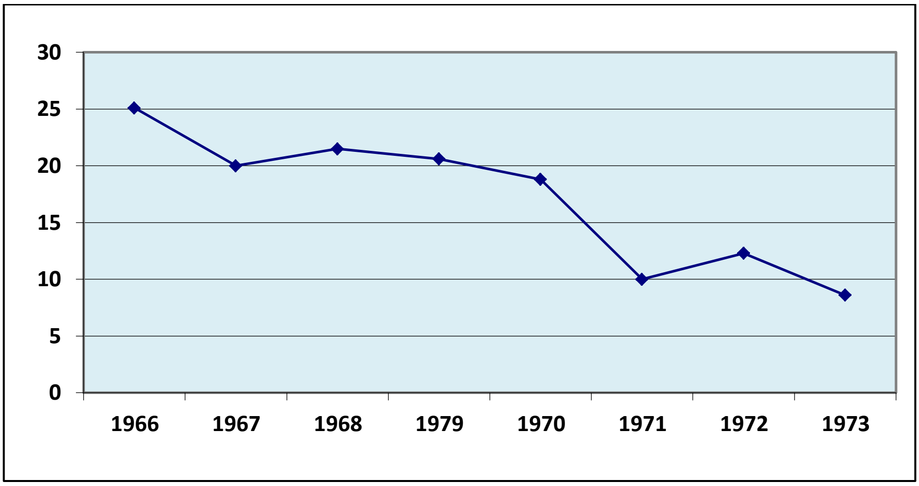
أولاً:- الاتجاه العام: توجد مؤثرات تجعل الظاهرة تتبع مجرى ثابت عبر فترة طويلة من الزمن. فبالرغم من وجود تعرجات في المنحنى التاريخي للمسلسلة تتسببها التغيرات الموسمية والدورية والعشوائية نجد أن هناك اتجاه خاص تتجه الظاهرة على مدى زمني طويل. هذا الاتجاه قد يكون تصاعديا كزيادة ثاني أكسيد الكربون الذي بدأ في حدود 280 جزء بالمليون قبل عام 1880 وهو في زيادة مستمرة نتيجة زيادة حرق الوقود الأحفوري لتوليد الطاقة وهو يقترب من حدود 550 جزء بالمليون في الوقت الحالي ويتسبب في ظاهرة الاحتباس الحراري العالمي وتغير المناخ. وقد يكون الاتجاه العام في التناقص كتناقص إنتاج النفط من الحقول نتيجة تناقص المخزون، وتناقص مساحات الغابات في العالم جراء الزحف العمراني والزراعي والتصحر، والشكل التالي يبين الاتجاه العام لصادرات ميناء طرابلس هو التناقص نتيجة الاعتماد على النفط كسلعة أساسية للتصدير.

جدول(57) صادرات ميناء طرابلس للفترة 1966–1973 بآلاف لأطنان.

السنة	66	67	68	69	70	71	72	73
الكمية	25.1	20.0	21.5	20.6	18.9	10.0	12.3	8.6

المصدر: أبوالقاسم العزابي وفوزي الأسدي و أبو صفحة وبشير أبوقيلة، دليل الباحث، ط 2، المنشأة العامة للنشر التوزيع والإعلان، طرابلس، 1982.

شكل(12) صادرات ميناء طرابلس للفترة 1966 –1973 (ألف طن).



ثانياً:– التغيرات الموسمية: هذه التقلبات تطرأ على الظاهرة بشكل منتظم في مواسم معينة من السنة، كزيادة مبيعات الأغنام في أعياد الأضاحي، وزيادة مبيعات الملابس في عيد الفطر، والبرودة في الشتاء والحرارة في الصيف والمد العالي البحري والجزر المرتبطة بدوران القمر حول الأرض.

ثالثاً:– التغيرات الدورية في فترات أطول من السنة: وهي تحدث على مدى أطول من المتغيرات الموسمية مثل دورات البقع الشمسية لفترات 11 و 22 سنة المعروفة، ودورات الزلازل والبراكين ودورات الجفاف. والتغيرات الموسمية والدورية تحدثان متزامنتان مع خط الاتجاه العام .

رابعاً:– التغيرات العشوائية Random oscillations: وهذه تحدث بسبب ظروف خارجية فجائية لا يمكن التنبؤ بها فهي لا تحدث طبقا لقاعدة معينة بل لعوامل فجائية كالحروب وانهيارات السدود أو الفيضانات النهرية.

الاتجاه العام:

الاتجاه العام يعني تغير في المتوسط العام لفترات متلاحقة إما بالزيادة أو التناقص التدريجي نحو وضع جديد يختلف عن سابقه. وقد تكون فترة الاتجاه قصيرة جدا لتمثل جزءا من دورة أو

قد يستمر هذا الاتجاه لفترة طويلة قد تمتد إلى مئات السنين، ولدراسة التغيرات المناخية، على سبيل المثال، يحتاج الباحث إلى تحليل مسلسلات زمنية طويلة ولا يكتفي بمسلسلات قصيرة قد يستنتج منها معلومات غير صحيحة.

ولأجل التعرف على اتجاهات المناخ يستخدم أسلوب Semi- average data حيث تقسم المسلسلة الزمنية إلى نصفين ويحسب المتوسط الحسابي لكل فترة على حده وعلى تخطيط الانتشار للمسلسلة ككل نوقع قيمة المتوسط الأول و قيمة المتوسط الثاني ثم نوصل النقطتين بخط مستقيم يمر بوسط نقاط التوزيع الأخرى ولا يلتفت إليها، عندها نحصل على خط الاتجاه الذي يبين فيما إذا كان المناخ متجها صعودا أو هبوطا. ويمكن اختبار الدلالة الإحصائية لهذا التغير، باستخدام نموذج الفرق بين المتوسطين كما يلي:

مثال: الجدول التالي (58) يحتوي على قيم المطر بمحطة أرصاد مطار طرابلس الدولي خلال الفترة الممتدة من السنة المطرية 60/1959 إلى 85/1984 والمطلوب اختبار هل هناك فرق دال إحصائيا بنسبة ثقة 95% بين متوسطي الفترتين الأولى الممتدة من 60/1959 إلى 72/1971 والفترة الثانية الممتدة من 73/1972 إلى 85/1984.

جدول (58) قيم المطر المسجلة بمطار طرابلس الدولي خلال الفترة الممتدة من السنة المطرية 60/1959 إلى 85/1984

الفترة الثانية		الفترة الأولى	
المطر ملم	السنة	المطر ملم	السنة
365.6	73-72	189.6	60-59
325.2	74-73	285.5	61-60
304.4	75-74	320.6	62-61
501.5	76-75	365.2	63-62
224.4	77-76	474.2	64-63
351.8	78-77	252.5	65-64
356.7	79-78	261.2	66-65
182.6	80-79	338.2	67-66
516.7	81-80	198.7	68-67
292.1	82-81	169.0	69-68
244.3	83-82	164.0	70-69
345.3	84-83	217.0	71-70
387.8	85-84	366.8	72-71

من الجدول يتبين أن المعطيات هي كما يلي:

الفترة الأولى	الفترة الثانية
M=13	N=13
$\sum X_1=3235$	$\sum X_2=4010.6$
$\bar{X}_1 = 269.64$	$\bar{X}_2 = 334.2$
$S_1=92.6$	$S_2=99.6$

خطوات الاختبار :

أولاً:- الفرضيات : الفرضية العدمية تنص على أنه ليس هناك فرق بين المتوسطين مما يعني عدم وجود تغير مناخي، أما الفرضية البديلة فهي تقول بوجود تغير مناخي بدليل أن المتوسطين غير متساويين. رياضيا تصاغ الفرضيتان بالشكل التالي:

$$H0: \overline{X}_1 = \overline{X}_2$$

$$HA: \overline{X}_1 \# \overline{X}_1$$

ثانياً:- الاختبار المناسب هو التائي ويحسب بالمعادلة التالية:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{Sp\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

على اعتبار أن:

m و n هما طول الفترة الأولى وطول الفترة الثانية على التوالي.

Sp = الانحراف المعياري المرجح من العينتين ويقدر بالمعادلة التالية :

$$Sp = \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}}$$

أما S_1 و S_2 فهما الانحراف المعياري للفترة الأولى والانحراف المعياري للفترة الثانية على التوالي.

ثالثاً:- المقارنة والاستنتاج :

من الجدول التائي نقدر قيمة T المعيارية عند α 0.05 ودرجات الحرية df التي تساوي حاصل جمع العينتين مطروح منه 2 أي : (m+n-2) .

قاعدة الاختبار: ارفض H0 إذا كانت القيمة المطلقة ل T الاختبارية أكبر من قيمة T الجدولية واستنتج بأن التغير المناخي في هذه الحالة دال إحصائيا وليس ناتجا عن الصدفة، أما إذا حصل العكس وكانت قيمة T الاختبارية أصغر من T الجدولية فاقبل الفرضية الصفرية واستنتج بأن الفرق بين المتوسطين صغير ولا يمكن اعتباره مختلفا جذريا عن الصفر، بما يعني عدم وجود تغير مناخي ذو دلالة إحصائية.

ثالثاً:- حساب قيمة الانحراف المعياري المرجح ثم قيمة الاختبار التالي:

$$Sp = \sqrt{\frac{(13-1)92.6^2 + (13-1)99.6^2}{13+13-2}} = 96.16$$

$$T = \frac{269.64 - 334.2}{96.16\sqrt{\frac{1}{13} + \frac{1}{13}}} = -1.71$$

رابعاً:- قيمة T الجدولية بنسبة ثقة 95% ودرجات حرية 24 تساوي 2.064

خامساً:- المقارنة والاستنتاج: بما أن t الاختبارية أصغر من T الجدولية نقبل الفرضية الصفرية ونستنتج أن الفرق بين المتوسطين ليس كبير بحيث يجعلنا نجزم بحصول تغير مناخي دال إحصائيا.

تحليل التطرف المطري بمحطة أرصاد نالوت:

التحليل التالي يستخدم سجل أكبر كمية مطر يومي سقطت في السنة خلال الفترة الممتدة من سنة 1945 إلى سنة 1996 بمحطة أرصاد نالوت على جبل نفوسة لإثبات فرضية التغير المناخي أو رفضها. لأجل ذلك قسمت فترة السجل إلى فترتين الأولى من عام 1945 إلى عام 1970 والثانية من عام 1971 إلى عام 1996 لاختبار الفرضية القائلة بأن متوسط كمية التطرف المطري للفترة الثانية أكبر من متوسط الفترة الأولى تمشيا مع التغير المناخي الذي تسببه ظاهرة الاحتباس الحراري الذي تظهر آثاره في زيادة تكرار وشدة التطرفات المطرية ولإتمام الاختبار نستخدم النموذج الإحصائي التالي:

جدول(59) أكبر كمية مطر سقطت في يوم(ملم) بمحطة أرصاد مدينة نالوت

خلال الفترة الأولى الممتدة من 1945 – 1970

27.8	25.6	13.4	16.8	45.6	29.2	19.7	11.7	43.8
89.6	49.0	26.2	14.2	28.6	39.9	20.5	69.5	51.1
	29.3	14.4	11.6	33.5	19.0	32.7	16.0	30.2

معطيات الفترة الأولى:

مجموع المطر خلال الفترة الأولى = 808.9 ملم

متوسط المطر للفترة الأولى $\bar{X}_1$ = 31.11 ملم

المجموع التربيعي للفرق بين قيم المطر والمتوسط $\sum(X-\bar{X}_1)^2$ = 8859.49ملم

الخطوة التالية هي حساب قيمة الانحراف المعياري للفترة الأولى بالمعادلة التالية :

$$S_1=\sqrt{\frac{8859.49}{26-1}}=18.82$$

جدول(60) أكبر كمية مطر سقطت في يوم (ملم) بمحطة أرصاد مدينة نالوت

خلال الفترة الممتدة من 1971– 1996.

35.7	18.0	58.4	69.0	30.3	102.2	10.7	84.0	21.6
28.0	55.0	50.0	57.7	28.0	30.5	58.0	7.0	35.2
	51.5	125.2	36.5	45.0	47.0	43.0	38.0	16.2

معطيات الفترة الثانية:

مجموع المطر خلال الفترة الثانية = 1181.5 ملم

متوسط المطر للفترة الثانية $\bar{X}_2$ = 45.44 ملم

المجموع التربيعي للفرق بين قيم المطر والمتوسط $\sum(X-\bar{X}_2)^2$ = 11645.31 ملم

الخطوة التالية هي حساب قيمة الانحراف المعياري للفترة الثانية:

$$S_2=\sqrt{\frac{11645.31}{26-1}}=21.58$$

ويتطلب التحليل حساب قيمة الانحراف المعياري المرجح من الفترتين بالمعادلة التالية:

$$S_P=\sqrt{\frac{(M-1)S_1^2+(N-1)S_2^2}{M+N-2}}$$

$$2Sp=\sqrt{\frac{(26-1)18.82^2+(26-1)21.58^2}{52-2}}=24.20$$

الآن توفر لنا كل المعطيات الضرورية لحساب قيمة الاختبار التائي بالمعادلة التالية:

$$T=\sqrt{\frac{45.44-31.11}{20.24\sqrt{\frac{1}{26}+\frac{1}{26}}}}=2.55$$

الفرضيات: من مجرد مقارنة قيمتي المتوسطين يتبين أن المتوسط الثاني (45.44 ملم) أكبر من المتوسط الأول بكثير (31.11ملم) ويقترح تطبيق اختبار الطرف الأيمن فتكون فرضية العدم H0 هي لا فرق بين متوسطي الفترتين بينما تقول البديلة HA أن متوسط الفترة الثانية أكبر. رياضيا تصاغ الفرضيتان بالصيغة التالية:

$$H0 : \bar{X}_2 = \bar{X}_1$$
$$HA: \bar{X}_2 > \bar{X}_1$$

الخطوة التالية هي تقدير قيمة T_a المعيارية من الجدول التائي لاختبار الطرف الواحد ومنه يتبين أن قيمة T_a المقدرة بدرجات حرية 50 ومعدل ثقة 95% تساوي1.67

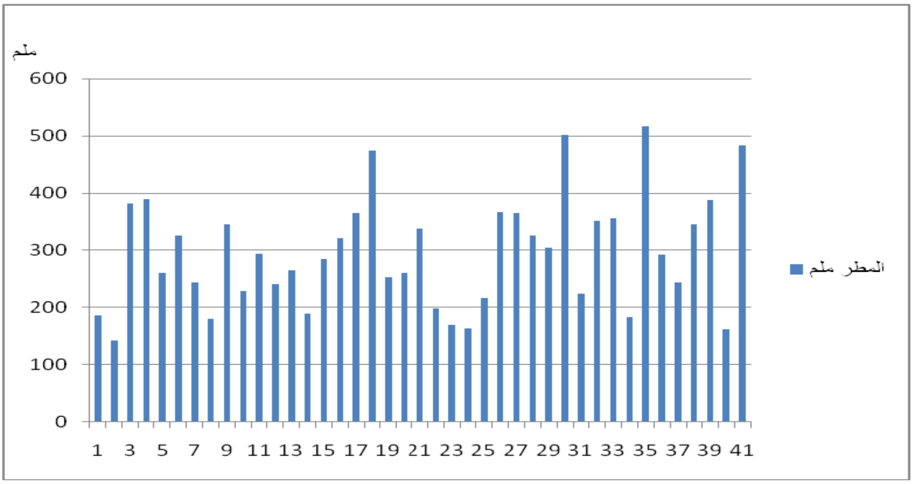
المقارنة والاستنتاج:

بما أن قيمة T الاختبارية (2.55) أكبر من قيمة T_a المعيارية (1.67) نرفض H0 ونقبل الفرضية البديلة التي تنص على أن المتوسط الثاني أكبر بدلالة (0.05).

نموذج الخمس سنوات المتدحرجة:

من خلال الرسم البياني لقيم المطر السنوي (الشكل 13) يتبين أن هناك تذبذب سنوي كبير في قيم المطر حول المتوسط الحسابي(291ملم) مما يشوش على الباحث ويصعب عليه اكتشاف الدورات في المناخ، لذلك يلجأ إلى أسلوب المتوسطات المتدحرجة الذي يقوم بطمس الذبذبات القصيرة والإبقاء على الدورات الطويلة فيسهل اكتشافها. وطول الفترة التي يجب استخدامها في حساب المتوسطات المتدحرجة يعتمد على الغرض من الدراسة وعلى طول مسلسلة البيانات المتوفرة. فإذا كان الغرض من الدراسة هو البحث عن التغيرات الطويلة الأمد يجب استخدام مسلسلة مناخية طويلة وكذلك متوسطات متدحرجة طويلة أيضا 21 أو 31 سنة، وإذا كان الاهتمام بدراسة الدورات والاتجاهات القصيرة فيمكن استخدام 3 أو 5 أو 7 سنوات متدحرجة.

شكل(13) كميات المطر المسجلة بمحطة أرصاد مطار طرابلس للفترة 49/1948 – 85/1984



يحسب متوسط كل فترة زمنيه طولها 5 سنوات بقسمة مجموع القيم للخمس سنوات على 5 كما يلي:

المتوسط الأول= مجموع قيم (س1+ س2 + س3 + س4 + س5) / 5 ويخصص الناتج للسنة (3) لأن هناك سنتين قبلها وسنتين بعدها، وفي الخطوة الثانية نحذف قيمة السنة الأولى ونضيف قيمة السنة السادسة للمجموع ثم نقسم على خمسة ونخصص الناتج للسنة الرابعة، وهكذا في كل مرة نحذف قيمة سنة من المقدمة ونضيف قيمة سنة حسب التسلسل. وفي نهاية المطاف نحصل

على مسلسلة جديدة فقدت سنتين من المقدمة وسنتين من المؤخرة، إذ لا يمكن حساب متوسط للسنة الأولى ولا الثانية لأنه لا يوجد قيم لسنوات قبلهما يمكن استعمالها في حساب المتوسط لهما، وبالمثل لا يمكن حساب متوسط للسنة الأخير ولا للتي قبلها لعدم توفر بيانات.

جدول (61) قيم المطر المسجلة بمطار طرابلس الدولي خلال الفترة الممتدة من السنة المطرية 1946 /47 إلى 87/1986

السنة	المطر ملم	السنة	المطر ملم	السنة	المطر ملم
47-1946	186.6	61-60	285.5	75-74	304.4
48-47	142.5	62-61	320.6	76-75	501.5
49-48	382.1	63-62	365.2	77-76	224.4
50-49	389.0	64-63	474.2	78-77	351.8
51-50	260.2	65-64	252.5	79-78	356.7
52-51	326.5	66-65	261.2	80-79	182.6
53-52	244.6	67-66	338.2	81-80	516.7
54-53	180.4	68-67	198.7	82-81	292.1
55-54	345.8	69-68	169.0	83-82	244.3
56-55	228.9	70-69	164.0	84-83	345.3
57-56	293.4	71-70	217.0	85-84	387.8
58-57	240.8	72-71	366.8	86-85	161.6
59-58	265.1	73-72	365.6	87-86	483.5
60-59	189.6	74-73	325.2		

الحل: نكون جدول جديد (62) يحتوي على مسلسلة البيانات الأصلية ومسلسلة الخمس سنوات المتدحرجة والعمود الأخير (على اليمين) يضع الرمز A للقيم الأكبر من المتوسط الحسابي 291.2 ملم ، والرمزB للقيم الأصغر من 291.2 ملم كما يلي:

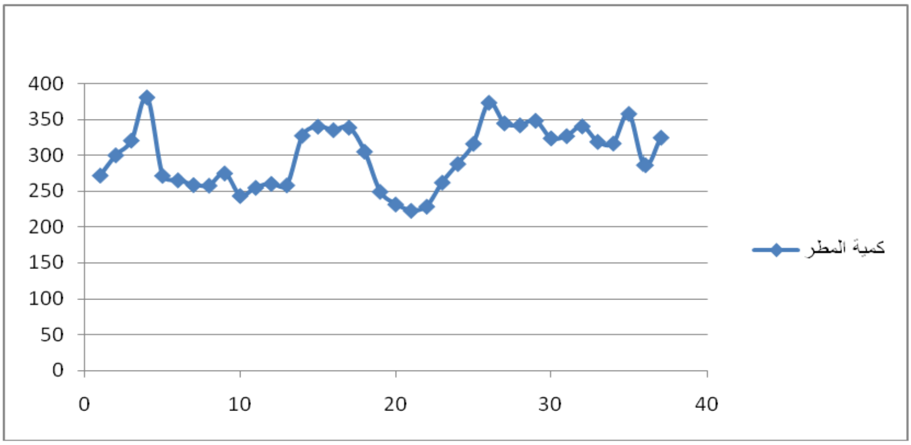
من الجدول (63) يتبين أن هناك فترتي جفاف بمتوسط طول 7.5 سنة، وثلاثة فترات مطيرة بمتوسط طول 6.3 سنة. ولاحظ أيضا أن هناك تباين كبير في الانحراف عن المتوسط يظهر في الفترات المطيرة أما متوسط الفترات الجافة فهو 7.5 سنة وهو أكثر تمثيلا لطول الفترتين 9 و 6 سنوات. وعلى ذلك يمكن استخدام متوسط فترات الجفاف كوسيلة للتنبؤ بطول فترة الجفاف القادمة التي من المتوقع لن تحيد كثيرا عن المتوسط، بينما لا يمكن استخدام متوسط الفترات المطيرة لأنه محسوب من عينة غير متجانسة يتراوح متوسطاتها بين 3 و11 سنة. وللحصول على أفضل النتائج يجب حساب المتوسط الحسابي من بيانات سجل طويل ولا يكون مقتصرا على عدد قليل من المفردات كما في المثال الذي تم سرده لتوضيح الفكرة فقط.

جدول (62) حساب مسلسلة الخمس سنوات المتدحرجة.

السنة المطرية	المطر ملم	الخمس سنوات المتدحرجة لقيم المطر	أكبر من المتوسط أو أصغر منه
47-1946	186.6		
48-47	142.5		
49-48	382.1	272.08	B
50-49	389.0	300.06	A
51-50	260.2	320.48	A
52-51	326.5	380.34	A
53-52	244.6	271.70	B
54-53	180.4	265.44	B
55-54	345.8	258.82	B
56-55	228.9	258.06	B
57-56	293.4	275.00	B
58-57	240.8	243.76	B
59-58	265.1	255.08	B
60-59	189.6	260.52	B
61-60	285.5	258.40	B
62-61	320.6	327.22	A
63-62	365.2	339.80	A
64-63	474.2	334.94	A
65-64	252.5	338.46	A
66-65	261.2	305.16	A
67-66	338.2	249.52	B
68-67	198.7	232.00	B
69-68	169.0	223.16	B
70-69	164.0	228.88	B
71-70	217.0	262.26	B
72-71	366.8	288.10	B
73-72	365.6	316.00	A
74-73	325.2	372.90	A
75-74	304.4	344.42	A
76-75	501.5	341.66	A
77-76	224.4	347.96	A
78-77	351.8	323.6	A
79-78	356.7	326.64	A
80-79	182.6	340.18	A
81-80	516.7	318.74	A
82-81	292.1	316.46	A
83-82	244.3	357.50	A
84-83	345.3	286.48	B
85-84	387.8	324.50	A
86-85	161.6		
87-86	483.5		

المتوسط السنوي للمطر المحسوب للفترة= 291.2 ملم

شكل(14) دورات الجفاف والمطر كما تظهرها مسلسلة الخمس سنوات المتدحرجة للمطر بمحطة أرصاد مطار طرابلس للفترة 1948/49 –1984/85



جدول (63) الملخص الإحصائي للفترات المطيرة والفترات الجافة المحسوبة لمسلسلة المطر بمحطة أرصاد مطار طرابلس للفترة الممتدة من السنة المطرية 1946/47 إلى 1986/87

الفترات الجافة	طول الفترة الجافة السنوات	الفترات المطيرة	طول الفترة المطيرة بالسنوات
61/60 – 53/52	9	52/51 –49/48	3
72/71 – 67/66	6	66/65 –62/61	5
		86/85 –73/72	11
متوسط الفترات الجافة	**7.5 سنوات**	**متوسط الفترات المطيرة**	**6.3 سنوات**

تمرين: الجدول (64) يعرض سجل المطر بمحطة أرصاد العجيلات للفترة 1960-1994(ملم)، والمطلوب تطوير الجدول لحساب مسلسلة الخمس سنوات المتدحرجة ورسم المسلسلة على ورقة رسم بياني لتوضيح الصورة ومعرفة الدورات، استرشد بالتوجيهات المبينة بأول المثال السابق.

جدول(64) مسلسلة أمطار مدينة العجيلات للفترة 1960-1993(ملم).

150.5	238.8	191.7	230.3	150.6
87.0	72.0	295.5	163.5	150.0
197.5	477.0	251.7	227.9	181.9
207.5	196.1	263.5	143.5	350.3
153.0	405.5	199.4	364.9	90.5
200.1	109.3	241.3	176.5	306.0
	112.7	166.4	54.7	128.5

البحث عن الدورات المناخية باستخدام نموذج باينوميال :

إن الذبذبات ذوات التردد العالي (الذبذبات القصيرة) تشوش على الدورات الكبيرة والاتجاهات الهامة في مسلسلة بيانات المطر وتجعل عملية اكتشافها أمرا صعبا على الباحث، لذا ينصح دائما بتصفيتها من مسلسلة البيانات عن طريق نموذج باينوميال binomial coefficient الوارد في البحث المنشور لمنظمة الأرصاد العالمية (WMO). وبتعبير آخر إن تطبيق نموذج باينوميال على مسلسلة البيانات يضمن إلغاء الذبذبات الأقصر من خمس سنوات ويبقي على الدورات والاتجاهات الطويلة في صورة واضحة يمكن التعرف عليها وقياسها بدقة، وللحصول على أفضل النتائج لا بد أن تكون مسلسلة البيانات طويلة بقدر الإمكان، ويحسب نموذج باينوميال بالقانون التالي:

$$b=0.02Xi_{-4}+0.05Xi_{-3}+0.12Xi_{-2}+0.20Xi_{-1}+0.22Xi+0.20Xi_{+1}+0.12Xi_{+2}+0.05Xi_{+3}+0.02Xi_{+4}$$

وبعد أن تصفى مسلسلة البيانات من الذبذبات القصيرة (العمودين الثالث والسادس في الجدول 65) تمثل بيانيا بتوقيع السنوات على المحور السيني وكميات المطر (قيم b من العمودين الثالث والسادس) على المحور الصادي. ومن تخطيط الانتشار للبيانات يمكن تحديد الفترات المطيرة وهي السنوات الواقعة فوق المتوسط الحسابي أما الفترات الجافة فهي الواقعة تحت المتوسط الحسابي ثم نكون جدول نحدد فيه فترات الجفاف والفترات المطيرة. لنأخذ بيانات مدينة العجيلات كمثال تطبيقي:

نكون جدول من ثلاثة أعمدة الأول للسنوات والثاني لكميات المطر والثالث نحسب فيه معامل باينوميال (كمية المطر الموزونة b) باستخدام المعادلة المبينة أعلاه كما يلي:

جدول (65) النتائج المتحصل عليها من تطبيق نموذج باينوميال على سجل المطر بالعجيلات.

معامل باينوميال b	كمية المطر ملم	السنة	معامل باينوميال b	كمية المطر ملم	السنة
224.0	263.5	1977		150.6	1960
209.1	196.1	1978		230.3	1961
210.7	207.5	1979		191.7	1962
221.2	90.5	1980		238.8	1963
240.6	364.9	1981	193.2	150.5	1964
260.7	199.4	1982	196.7	150.0	1965
262.5	405.5	1983	177.7	163.5	1966
255.3	153.0	1984	167.4	295.5	1967
237.5	306.0	1985	160.7	72.0	1968
216.6	176.5	1986	166.0	87.0	1969
194.2	241.3	1987	187.6	181.9	1970
173.3	109.3	1988	230.9	227.9	1971
152.6	200.1	1989	267.9	251.7	1972
	128.5	1990	287.3	477.0	1973
	54.7	1991	288.2	197.5	1974
	166.4	1992	265.1	350.3	1975
	112.7	1993	242.5	143.5	1976

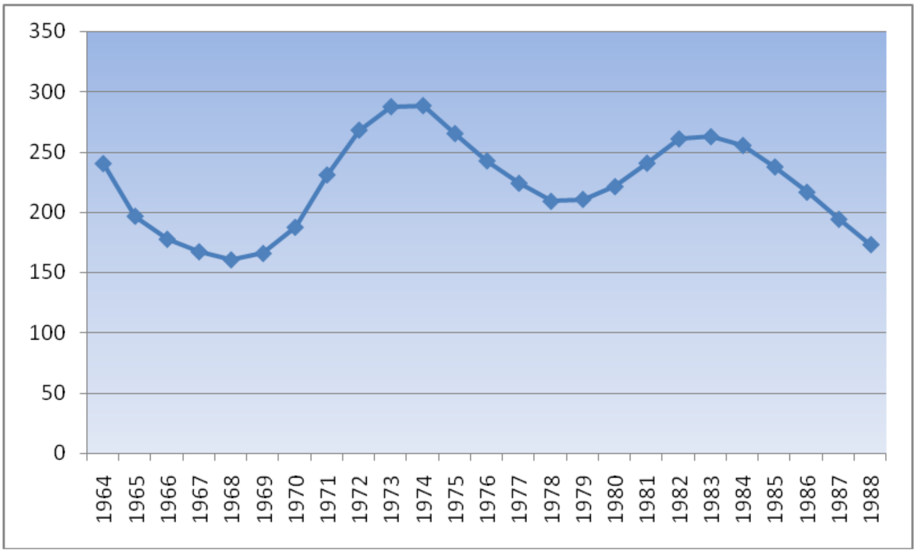
الآن نكون جدول جديد نقارن فيه قيم b بالمتوسط الحسابي المحسوب للفترة وهو (203.9ملم) ونعطي إشارة (+) إذا كانت b أكبر من (203.9ملم) وإشارة (−) إذا كانت أصغر من (203.9ملم) كما يلي:

جدول (66) مقارنة قيم باينوميال بالمتوسط الحسابي للفترة لتحديد سنوات الجفاف.

الإشارة	معامل باينوميال	السنة	الإشارة	معامل باينوميال	السنة
+	224.0	1977			1960
+	209.1	1978			1961
+	210.7	1979			1962
+	221.2	1980			1963
+	240.6	1981	+		1964
+	260.7	1982	-	196.7	1965
+	262.5	1983	-	177.7	1966
+	255.3	1984	-	167.4	1967
+	237.5	1985	-	160.7	1968
+	216.6	1986	-	166.0	1969
-	194.2	1987	-	187.6	1970
-	173.3	1988	+	230.9	1971
-	152.6	1989	+	267.9	1972
		1990	+	287.3	1973
		1991	+	288.2	1974
		1992	+	265.1	1975
		1993	+	242.5	1976

الخطوة التالية نوقع البيانات المتحصل عليها على ورقة رسم بياني عادية، السنوات على المحور الأفقي وكميات المطر الموزون b على المحور الرأسي ومنه يتبين الدورات المطيرة والجافة بدون التشويش الذي كانت تسببه الذبذبات القصيرة في البيانات.

شكل(15) دورات الجفاف والمطر كما تظهرها مسلسلة باينوميال.



أخيرا نكون جدول نحدد فيه الدورات المطيرة والدورات الجافة كما يلي:

جدول (67) دورات الجفاف ودورات المطر.

دورة مطيرة	دورة جفاف
1971- 1986 لمدة 16 سنه	1965- 1970 لمدة 6 سنوات
	1987-1989 لمدة 3 سنوات

وفيما يلي مثال لكيفية حساب معامل باينوميال بواسطة الحاسب اليدوي على سجل محطة العجيلات، بداية من سنة 1964 وهي السنة الأولى الممكن أن نحسب لها المعامل، وإذا أردت أن تحسب b لسنة 1963 على سبيل المثال، لن تتوفر البيانات إلا لثلاثة سنوات قبلها والنموذج يشترط أربعة سنوات وبالنسبة لسنة 1962 لا تتوفر بيانات إلا لسنتين قبلها أما سنة 1961 فالسنوات التي قبلها واحدة فقط. أما بالنسبة لعام 1972 لا يتوفر بيانات إلا لثلاثة سنوات والنموذج يشترط أربعة وهكذا. لاحظ أن تطبيق هذا النموذج يفقد المسلسلة الزمنية ثمانية سنوات أربعة من البداية وأربعة من الأخير لذلك ينصح دائما بأن تكون المسلسلة طويلة قدر الإمكان حتى تتحمل هذا الفاقد في المعلومات:

جدول (68) مثال لكيفية حساب معامل باينوميال

بواسطة الحاسب اليدوي

السنة	معامل الضرب									b
	.02	.05	.12	.20	.22	.20	.12	.05	.02	
					كمية المطر					
1960					150.6					
1961					230.3					
1962					191.7					
1963					283.8					
1964	150.6	230.3	191.7	283.8	150.5	150.0	163.5	295.5	72.0	193.2
1965	230.3	191.7	283.8	150.5	150.0	163.5	295.5	72.0	87.0	196.7
1966	191.7	283.8	150.5	150.0	163.5	295.5	72.0	87.0	181.9	177.7
1967	283.8	150.5	150.0	163.5	295.5	72.0	87.0	181.9	227.9	167.4
1968	150.5	150.0	163.5	295.5	72.0	87.0	181.9	227.9	251.7	160.7
1969	150.0	163.5	295.5	72.0	87.0	181.9	227.9	251.7	477.0	166.0
1970	163.5	295.5	72.0	87.0	181.9	227.9	251.7	477.0	197.0	187.6
1971	295.5	72.0	87.0	181.9	227.9	251.7	477.0	197.0	350.3	230.9
1972					251.7					
1973					477.0					
1974					197.0					
1975					350.3					

المصادر والمراجع:

- العزابي، أبو لقاسم محمد، (1995)، النقل والمواصلات، في كتاب الجماهيرية دراسة في الجغرافيا، الدار الجماهيرية للنشر والإعلان، سرت.
- أبو صالح محمد صبحي، (2000)، الطرق الإحصائية، دار اليازوري، عمان.
- أبو راضي، فتحي عبد العزيز، (1987)، مقدمة في الأساليب الكمية في الجغرافيا، دار المعرفة الجامعية، الإسكندرية.
- الأطرقجي، محمد على، (1980)، الوسائل التطبيقية في الطرق الإحصائية، دار الطليعة، بيروت.
- أبو عيانة، فتحي، (1987)، مدخل إلى التحليل الإحصائي في الجغرافيا البشرية، دار المعرفة الجامعية، الإسكندرية.
- الشريف، فوزي صالح، (2014)، المشكلات الاجتماعية لعدم الإنجاب على الأسرة، مجلة صدى المعرفة، العدد الأول، جمعية جبل نفوسه للتنمية، يونيو 2014.
- ثورة الفاتح في عشر سنوات. دار الشورى، بيروت، 1979.
- حميد، عريبي الشيباني، (2007)، التعليم العالي وتوظيفه في التنمية لخدمة المجتمع وتفعيل مراكز البحث العلمي، مجلة كلية الآداب، جامعة الفاتح.
- مجلة الطاقة والحياة، العدد (3)، يونيو 1994.
- مقيلي، إمحمد عياد، (2024)، الاحتباس الحراري العالمي وأثره على التطرف المطري والحراري بمحطتي أرصاد ليبية، مجلة ليبيا الدراسات الجغرافية، المجلد الرابع، العدد الثاني، يوليو 2024.
- مقيلي، إمحمد عياد، (1991)، اتجاهات الأمطار واحتمالات التصحر في منطقة الجفارة شمال غرب الجماهيرية الليبية، مجلة الدراسات الصحراوية، المجلد الأول، العدد الأول، مرزق.
- فياض، فتحي عبد الله، (1983)، مبادئ الإحصاء الجغرافي، المنشأة العامة للنشر والتوزيع والإعلان، طرابلس.
- عدس، عبد الرحمن، (1997)، مبادئ الإحصاء في التربية وعلم النفس، دار الفكر، عمان.
- شلبي، أحمد، (1989)، كيف تكتب بحثا أو رسالة، ط/2 ، مكتبة النهضة المصرية، القاهرة.

– شحادة، نعمان، (1997)، الأساليب الكمية في الجغرافيا باستخدام الحاسوب، دار صفاء، عمان.

– يونس، فضل أحمد، (1993)، الجغرافيا السياحية، دار النهضة العربية، بيروت.

- Christodolou, F. and A. d. Lewis. Distance Perception Of Reading Univ. Students. Univ. of Reading. Dept of Geography.1976.

- Maurice Yeates. An Introduction to Quantitative Analysis In Human Geography. Mc GRAW-HILL BOOK COMP. New .York. 1974.

- Khaznte, R.” Elementary Statistics in a world of application-Good Year pub. Comp., Califorenia.1979.

- Hanusehek ,E.A., and .J.E. Jackson.” Statistical Methods for social scientists. Academic Press, New York,1977.

- Neal, F. Jensen: Limits to Growth in World food Production. Science. Vol (201).28 July 1978.

-Critshfield, H. General Climatology, Prentice-Hall, Inc.; First Edition (January 1, 1960) P86.

- Gene V.Glass & Jullian C Stanley “ Statistical Methods in Education And Psychology” Prentice-Hall, Inc., New Jersey 1970.

- John Silk: Statistical Concepts in Geography. George Allen and Unwind. London. 1979.

-John Netter & Others. Applied Statistics. 1978.

- Living with Risk. A global review of disaster reduction nitiatives. Preliminary version ISDR, Geneva. July 2000.

-Tom, H.C.S.” Some Methods of Climatological Analysis.”WMO Technical Note”(81). 1966.

-Siegal, S. “Nonparametric Statistics for Behavioral Science.” McGraw Hill. 1957.

-Khazanie, R. “Elementary Statistics”. Good Year Publishing Comp. INC. Santa Monica, Cal. 1979.

-WMO, “Climatic Change.” WMO Technical Note, No (79) Geneva. 1966.

ملحق (1) الجدول الثائي

T

	t لاختبار الطرفين		t لاختبار الطرف الأوحد	
df	α =0.05	α =0.01	α =0.05	α =0.01
1	12.706	63.657	6.314	31.821
2	4.303	9.925	2.920	6.965
3	3.182	5.841	2.353	4.541
4	2.776	4.604	2.135	3.747
5	2.571	4.032	2.015	3.365
6	2.447	3.707	1.943	3.143
7	2.365	3.499	1.895	2.998
8	2.306	3.355	1.860	2.896
9	2.262	3.250	1.833	2.821
10	2.228	3.169	1.812	2.764
11	2.202	3.106	1.796	2.718
12	2.179	3.055	1.782	2.681
13	2.160	3.012	1.771	2.650
14	2.145	2.977	1.761	2.624
15	2.131	2.947	1.753	2.602
16	2.120	2.921	1.746	2.583
17	2.110	2.898	1.740	2.567
18	2.101	2.878	1.734	2.552
19	2.093	2.861	1.729	2.539
20	2.086	2.845	1.725	2.528
21	2.080	2.831	1.721	2.518
22	2.074	2.819	1.717	2.508
23	2.069	2.807	1.714	2.500
24	2.064	2.797	1.711	2.492
25	2.060	2.787	1.708	2.485
26	2.056	2.779	1.706	2.479
27	2.052	2.771	1.703	2.473
28	2.048	2.761	1.701	2.462
29	2.045	2.756	1.699	2.462
30	2.042	2.750	1.697	2.457
40	2.021	2.704	1.684	2.423
60	2.000	2.660	1.671	2.390
120	1.980	2.617	1.658	2.358
∞	1.960	2.576	1.645	2.326

إذا كانت قيمة الاختبار **t** على يمين القيمة المقدرة من الجدول أرفض **H0**

ملحق (2) جدول كاي تربيع

X^2

df	α =0.05	α =0.1	df	α =0.05	α =0.01
1	3.841	6.635	26	38.885	45.642
2	5.991	9.210	27	40.113	46.963
3	7.815	11.345	28	41.337	48.278
4	9.488	13.277	29	42.557	49.588
5	11.070	15.086	30	43.773	50.892
6	12.592	16.812	32	46.194	55.456
7	14.067	18.473	34	48.602	56.064
8	15.507	20.090	36	50.999	58.619
9	16.919	21.666	38	53.384	61.162
10	18.307	23.209	40	55.759	63.691
11	19.675	24.735	42	58.124	66.306
12	21.076	26.212	44	60.481	68.710
13	22.362	27.688	46	62.830	71.201
14	23.685	29.141	48	65.171	73.683
15	24.996	30.578	50	67.505	76.154
16	26.296	32.000	52	69.832	78.616
17	27.587	33.409	54	72.153	81.069
18	28.869	34.805	56	74.468	83.513
19	30.144	36.191	58	76.778	85.950
20	31.410	37.566	60	79.082	88.379
21	32.671	38.963	62	81.381	90.802
22	33.924	40.289	64	83.675	93.217
23	35.172	41.628	66	85.965	95.626
24	36.415	42.980	68	88.250	98.028
25	37.652	44.314	70	90.531	100.425

يتبع ملحق (2) جدول كاي تربيع X2

للطرفين(α=0.05)

	الأيمن	الأيسر		الأيمن	الأيسر
df	α =0.975	A=0.025	df	α =0.975	α =0. 025
1	5.024	0.001	22	36.781	10.982
2	7.378	0.051	24	39.364	12.401
3	9.348	0.216	26	41.923	13.844
4	11.143	0.484	28	44.461	15.408
5	12.833	0.831	30	46.979	16.791
6	14.449	1.237	40	59.342	24.433
7	16.013	1.690	50	71.420	32.357
8	17.535	2.180	60	83.298	40.482
9	19.023	2.700	70	95.023	48.758
10	20.483	3.247	80	106.629	57.153
11	21.920	3.816	90	118.136	65.647
12	23.337	4.404	100	129.562	74.222
13	24.736	5.009	110	140.917	82.867
14	26.119	5.629	120	152.211	91.573
15	27.488	6.262	130	163.453	100.331
16	28.845	6.908	140	174.648	109.137
17	30.191	7.564	150	185.800	117.985
18	31.526	8.231	200	241.058	162.728
19	32.852	8.907	500	563.852	439.936
20	34.170	9.591			

المصدر: محمد صبحي أبوصالح، وعدنان محمد عوض، مقدمة في الإحصاء، مركز بحوث العلوم الأساسية، طرابلس، 1983، ص ص 305-306.

يتبع ملحق (2) جدول كاي تربيع X^2
للطرفين (α=0.01)

df	α =0.975	A=0.025	df	α =0.975	α =0. 025
1	7.879	-	22	42.796	9.542
2	10.597	0.020	24	45.559	10.856
3	12.838	0.115	26	48.290	12.198
4	14.860	0.297	28	50.993	13.565
5	16.750	0.554	30	53.672	14.954
6	18.548	0.872	40	66.766	22.164
7	20.278	1.239	50	79.490	29.707
8	21.955	1.646	60	91.952	37.485
9	23.589	2.088	70	104.215	45.442
10	25.188	2.558	80	116.321	53.540
11	26.757	3.053	90	128.299	61.754
12	28.299	3.571	100	140.169	70.065
13	29.819	4.107	110	151.948	78.458
14	31.319	4.660	120	163.648	86.923
15	32.801	5.229	130	175.278	95.451
16	34.267	5.812	140	186.847	104.034
17	35.718	6.408	150	198.360	112.669
18	37.156	7.015	200	255.264	156.432
19	38.582	7.633	500	585.207	429.388
20	39.997	8.260			

المصدر: محمد صبحي أبوصالح، وعدنان محمد عوض، مقدمة في الإحصاء، مركز بحوث العلوم الأساسية، طرابلس، 1983، ص ص 305-306.

ملحق (3) الجدول الزاي Z-Score

Z لاختبار الطرفين		Z لاختبار الطرف الواحد	
50α =0.	α =0.01	α =0.05	α =0.01
1.96	2.58	1.65	2.33

ملحق (4) جدول التوزيع الطبيعي Z

$$Z = \frac{X - \overline{X}}{S}$$

Z	.00	.01	.02	.03	.04	.05	.06	.07	..08	.09
0.0	.0000	.0044	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0476	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2158	.2190	.2224
0.6	.2257	.2291	.2324	.2356	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4418	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4644	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.3817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.3964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4967	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.3993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.3996	.4997

المصدر: محمد صبحي أبو صالح وعدنان محمد عوض، **مقدمة في الإحصاء**، مركز بحوث العلوم الأساسية، طرابلس، 1983، ص302. نقلا عن هويل وجيسين:

Hoel and Jessen, Basic Statistics for Business and Economics. 2nd ed(1977). John Wiley& Sons.

الاستاذ الدكتور إمحمد عياد محمد مقيلي

- مواليد: سوق الجمعة طرابلس 1949/1/1
- تلقى تعليمه للمرحلتين الابتدائية والاعدادية والثانوية بمدرسة الجديدة، سوق الجمعة، طرابلس.
- الليسانس في الجغرافيا، كلية التربية، جامعة طرابلس، 1974م.
- الماجستير في الجغرافيا، جامعة إنديانا، الولايات المتحدة الأمريكية، 1979م.
- الدكتوراه في الجغرافيا، جامعة إنديانا، الولايات المتحدة الأمريكية، 1984م.
- العمل استاذ شرف بقسم الجغرافيا، كلية الآداب، جامعة طرابلس.

الكتب المؤلفة:ـ

- كتاب تلوث البيئة الطبيعية(مشترك)، منشورات الجامعة المفتوحة، طرابلس، ليبيا، 1990م.
- كتاب مدخل إلى البحث الجغرافي (مشترك)، دار الفلاح الكويت، 2000.
- كتاب مقدمة في الطقس والمناخ، منشورات الجامعة المفتوحة، طرابلس، ليبيا، 1993م.
- كتاب التلوث البيئي، دار شموع الثقافة، الزاوية، ليبيا، 2000م.
- مخاطر الجفاف والتصحر والظواهر المصاحبة لهما، دار شموع الثقافة، الزاوية، ليبيا، 2003م.
- المخاطر الهيدروجيموفولوجية، دار شموع الثقافة، الزاوية، ليبيا، 2003م.
- تطرفات الطقس والمناخ، دار شموع الثقافة، الزاوية، ليبيا، 2003م.

منشورات مركز البحوث والاستشارات - جامعة سرت
الطبعة الأولى 2025