جامعة الفرات

كلية الهندسة الزراعية بالحسكة

ع + م

٢٢٠

٢٦/٢٥

مقرر "الإحصاء وتصميم التجارب"

لطلبة السنة الثانية

الجزء النظري + الجزء العملي

إعداد :

أ.د. قصي العمر

د. إبراهيم البرهو

العام الدراسي ٢٠٢٥ / ٢٠٢٦

الفصل الأول

إعداد : د . قصي العمر

مدخل إلى علم الإحصاء

مقدمة : Introduction

ورد ذكر الإحصاء في القرآن الكريم في مواقع عدة " **وإن تعدو نعمة الله لا تحصوها** " ، " **لقد أحصاهم وعدهم عدا** " . " **ووضع الكتاب فترى المجرمين مشفقين مما فيه ويقولون يا ويلتنا مال هذا الكتاب لا يغادر صغيرة ولا كبيرة إلا أحصاها** " .... وغيرها من الآيات الكريمة . وقد ظهر الإحصاء في العصور القديمة مبنياً على فكرة **الحصر والعد** ، وأول من استخدمه الفراعنة حيث ساعدهم حصر وعد السكان والثروات في بناء الأهرامات ، كما استخدمت هذه الفكرة في ظل الدولة الإسلامية من قبل الخليفة المأمون لتحديد نسبة الزكاة على المسلمين .

وقد نشأت النظم الإحصائية مع نشوء الدولة ووجودها على وجه الأرض ، فمن أبسط الأمور أن أي حكومة تحتاج إلى معرفة المساحات المزروعة وإنتاجيتها من المحاصيل الإستراتيجية كالقمح والقطن ، والقادرين على حمل السلاح من السكان ، وعدد السكان القادرين على دفع الضرائب التي تفرض عليهم وذلك لكي تتمكن من إدارة دفة البلاد . وكذلك تهتم الدول المتقدمة بمعرفة خريطة توزيع القدرات العقلية والذهنية بين أفراد شعبها ، ليتم من خلال هذا المسح العام توزيع الطلاب على التعليم المناسب لهم ، ووضع كل فرد في المهنة والعمل المناسب لتفكيره وميوله ، ويتم تجنيد الشباب البالغين كل منهم في السلاح المناسب لقدراته ومواهبه .

وقد تطورت النظم الإحصائية من مفهوم الحصر والعد ولغة الأرقام إلى مفهوم الإحصاء وبمعناه العلمي، والذي أصبح يعد الآن من أهم أدوات البحث العلمي، حيث يهتم بالدراسة الكمية للظواهر المختلفة من اجل وصفها وتحليلها وتفسيرها والتنبؤ بها ثم التحكم فيها.

تعريف الإحصاء : Definition of Statistics

يشير معنى كلمة إحصاء إلى لفظ لاتيني يدعى "ستاتو" والذي يعني الدولة ، ولذلك فإن كلمة إحصاء في اللغة الانكليزية تعني STATISTICS والمشتقة من كلمة STATE وهي الدولة، وعليه فإن من أوائل التعريفات التي وضعت للإحصاء هو أنه : " **مجموعة الحقائق الخاصة بشؤون الدولة** " ، ثم تطور هذا المفهوم شيئاً فشيئاً حتى غدا من أهم أدوات البحث العلمي كما أسلفنا ، وبالتالي فإن الإحصاء أصبح علماً قائماً بحد ذاته له طرقه وأدواته وأساليبه ، كما أنه

أصبح من العلوم الضرورية للعلوم الأخرى كالعلوم الصناعية والتجارية والهندسية والزراعية والإنسانية والاقتصادية وغيرها.

ويعرف الإحصاء كعلم بأنه :

" مجموعة الطرائق والأدوات العلمية التي تتبع في الحصول على البيانات الكمية والنوعية ، وتصنيفها وتبويبها وعرضها وتحليلها و استخلاص النتائج منها".

أنواع الإحصاء : Types of statistics

للإحصاء نوعان أو وظيفتان أساسيتان : وهما :

١-الإحصاء الوصفي: Descriptive statistics

وهو ذلك الجزء من الإحصاء الذي يصف الواقع كما هو عليه الأمر في الحقيقة . فهو يعني عرض حقيقة الظاهرة المدروسة بشكل ملخص بالأساليب الإحصائية المتبعة في جمع البيانات وتبويبها وعرضها وتحليلها.

٢-الإحصاء الاستدلالي أو الاستقرائي: Inferential or Inductive statistics

وهو ذلك الجزء من الإحصاء الذي يهدف للتوصل إلى صورة للمجتمع الإحصائي عن طريق دراسة العينة ، فهو يهتم بالقيام باستدلالات حول المجتمعات الأصلية تستند على عينات مأخوذة منها. فغالباً يتم اللجوء إلى العينات في الكثير من الدراسات لأسباب عديدة ، ومن أهمها توفير المال والجهد والوقت المطلوب عند إجراء الدراسة عن طريق الحصر الشامل لوحدات المجتمع المدروس، بالإضافة إلى استحالة إجراء الكثير من الدراسات الإحصائية عن طريق الحصر الشامل كحالة دراسة نسبة ملوحة بحيرة أو نسبة التلوث الجوي فوق إحدى المدن بدخان عوادم السيارات أو عدم إمكانية فتح كل زجاجات الدواء المنتجة لمعرفة إن كان الدواء فيها محقق لشروط محددة أم لا.

مفاهيم إحصائية : Statistical concepts

نستعرض فيما يلي بعض المفاهيم الإحصائية التي لا بد من التعرف على مضمونها لكل دارس لعلم الإحصاء ، وهذه المفاهيم هي :

المتغير أو المتحول: Variable

وهو تلك الخاصية أو الكمية أو الصفة التي تتغير قيمتها من عنصر لآخر أو من مشاهدة إلى أخرى ، مثلاً : الطول متغير حيث تختلف هذه الخاصية من عنصر إلى عنصر أخر، وفي

العلوم الزراعية هناك متغيرات مثل : طول النبات ، نسبة الإنبات ، عدد الاشطاءات ، وزن البيضة ، نوع التربة ..... الخ. والمتغير يقابل مع الثابت Constant ، والثوابت هي السمات أو الخواص التي لا تتغير قيمتها أبدا، مثل العدد الذري للهدروجين ١، الكتلة الذرية أو الجزيئية للأوكسجين١٦ ، π = 3.14 ... الخ . والمتغير سواء أكان كمياً (الطول) أو نوعياً (نوع التربة) يظهر على نوعين، فإما أن يكون متصلاً أو منفصلاً.

١-المتغير المتصل أو المستمر: Continuous variable

هو ذلك المتغير الذي تختلف قيمه بمقادير صغيرة صغراً لا نهائياً، فالعمر مثلا هو متغير متصل لأنه لا يمكن أن نمر من عمر إلى عمر أخر مهما كان قريبا منه إلا إذا مررنا بعدد لا نهائي من الأعمار، ونحصل على قيمة هذا المتغير بالقياس لا بالعد ، كما أن الزمن متغير متصل ونحن نقيسه بواسطة الساعة وبشكل تقريبي وعلى درجة معينة من الدقة قد تصل إلى ۰.۰۰۱ من الثانية أو أكثر من ذلك ، ومنها : الأطوال، الأوزان، أجورعمال، درجات الحرارة ، درجات الاختبارات....الخ.

٢-المتغير المنفصل أو المتقطع أو الوثاب: Discrete variable

هو ذلك المتغير الذي تختلف قيمه بمقادير محددة ، وغالباً ما تكون القيم التي يأخذها ذلك المتغير أعداداً صحيحة كعدد البذور ، عدد أطفال عائلة ، عدد سكان مدينة .....الخ فالمتغير المنفصل نحصل على قيمه بالعد ، وهنا تقفز قيم المتغير من عدد صحيح إلى أخر متجاوزة ما بين العددين من أعداد كسرية ، فمثلاً لا يعقل أن يكون عدد البنين في فصل دراسي هو ٢٢,٥ أو ٢٨,٩ ، فالمتغير المنفصل يتميز عن المتصل بأنه يمكن عد عناصر مجال هذا المتغير ، كما توجد قفزات بين عناصره .

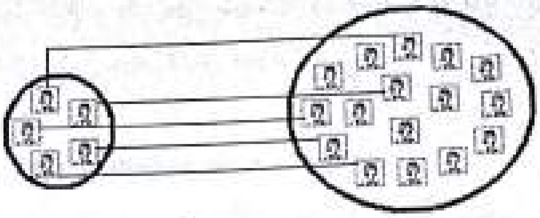
البيانات أو المعطيات (data) : هي الأرقام أو القياسات التي يأخذها المتغير ، والتي تجمع نتيجة لما نقوم به من ملاحظات .

المجتمع الإحصائي : Population

هو مجموعة كاملة من الأفراد أو الأشياء أو القياسات لها صفة مميزة عامة قابلة للملاحظة ، فإذا أردنا أن ندرس مدى ممارسة نساء كلية الزراعة للعمل بعد تخرجهن. فسوف يتألف المجتمع الإحصائي من مجموع خريجات كلية الزراعة .

العينة: Sample

وهي منظومة فرعية من المجتمع الاحصائي مثل خريجات كلية الزراعة في جامعة دمشق أو



الفرات فقط .

الثابت أو المعلمة أو المقياس الثابت (parameter) : هو أي صفة من صفات المجتمع الأصلي تكون قابلة للقياس وتتحدد قيمتها من جميع المشاهدات على كامل أفراد المجتمع ، مثل متوسط أعمار خريجات كلية الزراعة في الجامعات السورية .

المقياس الإحصائي أو الإحصاء (statistic) : وهو متغير تتحدد قيمته من العينة ، ويستخدم المقياس الإحصائي المحسوب من عينة لكي نقدر المقياس الثابت للمجتمع الأصلي ، مثل متوسط أعمار خريجات كلية الزراعة في جامعة دمشق أو الفرات .

مثال لمعرفة الفرق بين المصطلحات الإحصائية السابقة :

شركة تنتج براغي مكنات أو آلات ضمن ثخن معين يجب المحافظة عليه ، ولكي نحقق ونضبط هذا الثخن فواجب علينا أن نختار عدد من البراغي المنتجة خلال اليوم ونقيسها بعناية ودقة .

المتغير الذي نهتم بقياسه هو **ثخن البرغي** .

البيانات هي **قياسات** كل البراغي التي اختيرت في العينة .

المجتمع الأصلي الذي نهتم بالتعميم عليه هو **كل البراغي** المنتجة من الشركة خلال يوم كامل .

العينة هي البراغي المختارة .

المقياس الثابت هو **معدل ثخن كل البراغي** المنتجة لليوم كاملاً ، ويسمى معلمة أو بارا متر parameter .

المقياس الإحصائي هو **معدل ثخن البراغي في العينة** ، ويسمى statistic أو إحصاء .

ومن هذا المثال نلاحظ أنه من غير المحتمل أن تُعرف المعلمة على الإطلاق ، وذلك لصعوبة إجراء ذلك لأسباب اقتصادية وغير اقتصادية ، ولكن يمكن أن يُستدل عليها من المقاييس الإحصائية للعينة .

أساليب البحث الميداني :

يوجد أسلوبان ، وهما :

١-أسلوب الحصر الشامل :

يستخدم عندما يكون المجتمع الإحصائي غير كبير نسبياً، كما يتم اللجوء إليه أيضا على الرغم من ضخامة حجم المجتمع الإحصائي عند الضرورات، وهذا السلوك تقوم به الدولة عادة من خلال جهازها الإحصائي المؤهل والمتخصص في جمع البيانات، واهم مثال عليه هو التعدادات السكانية الشاملة .

ويتميز هذا الأسلوب بعدم إهماله لأي جزء من المجتمع الإحصائي المدروس، كما يتميز بدقة نتائج الدراسة. أما مساوئه فهي تتجسد في التكلفة المالية الكبيرة والجهد البالغ والزمن الطويل لإجرائه، هذا الزمن الطويل الذي قد يفقد البيانات أهميتها بسبب أنها تصف ما كان عليه المجتمع الإحصائي لسنوات عدة مضت وليس الوقت الحاضر.

٢-أسلوب العينة أو المعاينة :

ويتم اللجوء إليه للأسباب التالية:

أ- توفير الوقت والجهد والمال.

ب- عدم توفر الجهاز الإحصائي القادر على القيام بالدراسة في كافة مراحلها.

ب-استحالة القيام بالدراسة عن طريق المسح الشامل بسبب لا محدودية المجتمع الإحصائي المدروس مثل عدم إمكانية إجراء المسح الشامل في حال دراسة نسبة الملوحة في نهر الخابور.

ت-استحالة القيام بالدراسة عن طريق المسح الشامل بسبب عدم هدر أو إتلاف جميع وحدات المجتمع الإحصائي المدروس مثل عدم إمكانية إجراء المسح الشامل في حال دراسة نسبة البيض التالف من إجمالي إنتاج مدجنة.

ويتميز هذا الأسلوب بتوفير الوقت والجهد والمال، أما مساوئه فتنحصر في الانحرافات البسيطة عن النتائج والخصائص الحقيقية للمجتمع ، وهذا لا يبرر في كثير من الحالات اللجوء إلى أسلوب الحصر الشامل .

العينات : Samples

العينة هي نموذج يشمل جزء من فئات المجتمع الأصلي يكون ممثلا له تمثيلا جيدا، بحيث يحمل صفاته المشتركة ، وهذا النموذج أو الجزء يغني الباحث عن دراسة كل فئات ومفردات المجتمع الأصلي خاصة في حالة صعوبة واستحالة دراسة كل تلك الفئات . ويتم اختيار العينة وفقا لأسس وأساليب علمية متعارف عليها .

خطوات اختيار عينات البحث :

١- تحديد مجتمع البحث الأصلي تحديدا واضحا ودقيقا .

٢- تشخيص أفراد المجتمع أي إعداد قوائم بأسماء جميع الأفراد في المجتمع الأصلي للدراسة .

٣- اختيار وتحديد نوع العينة كأن تكون عينة طبقية تناسبية أو عينة منتظمة أو عينة عشوائية تعطى الفرصة لكل أفراد المجتمع الأصلي أن يكون ضمن العينة . وأساس تحديد نوع العينة هو أن العينة الجيدة والسليمة هي العينة التي تعكس خصائص المجتمع الأصلي وتمثله تمثيلا صحيحا ودقيقا .

٤- تحديد العدد المطلوب من الأفراد أو الوحدات في العينة، ويقصد به تحديد حجم العينة المختارة .

ويتأثر تحديد حجم العينة بعدة عوامل من أهمها :

أ) مدى التجانس أو التباين في خصائص المجتمع .

ب) مقدار الوقت المتوفر لدى الباحث وإمكاناته العلمية والمادية .

ج) درجة الدقة المطلوبة في البحث ومستوى الثقة .

أنواع العينات :

يمكن تحديد الأنواع المختلفة للعينات وفقاً لدرجة دقتها وتمثيلها للمجتمع الأصلي كالآتي:

١ - العينة الطبقية

٢- العينة الطبقية التناسبية

٣- العينة العشوائية البسيطة

٤- العينة العشوائية المنتظمة

٥- العينة العمدية أو الغرضية

٦- العينة العرضية أو عينة الصدفة

وسنوضح كل عينة من هذه العينات فيما يلي :

١- العينة العشوائية الطبقية: Stratified Random Sample

وفيها يقسم المجتمع إلى الشرائح أو الطبقات التي يشتمل عليها ، ومثال ذلك إذا كان مجتمع البحث يتكون من طلاب كلية الزراعة وحجم العينة المطلوبة للبحث هو ٤٠٠ طالب ، فيمكن أن تكون شرائح المجتمع وطبقاته مشكلة من الأقسام العلمية للكلية (قسم الأراضي، قسم المحاصيل، قسم الاقتصاد والإرشاد الزراعي ، قسم البيئة ، قسم الإنتاج الحيواني ----الخ) . فإذا كان عدد الأقسام خمسة يتم أخذ عدد (٨٠ طالب) من كل شريحة ، وإذا زاد عدد الأقسام عن الخمسة يقسم مجموع العينة المطلوبة عليها ثم يؤخذ عدد متساوي من كل منها ، فمثلا إذا كان عدد الأقسام ثمانية ، يؤخذ (٥٠) طالب من كل قسم (٤٠٠÷ ٨) وهكذا.

٢- العينة الطبقية التناسبية أو العينة الحصصية: Quota Sample

وفيها يتم تقسيم المجتمع الأصلي للبحث إلى شرائح أو فئات أو طبقات ، إلا أنه بدلا من تحديد حجم العينة على أساس متساوي من كل شريحة من شرائح المجتمع، فإننا نقوم بتحديد عدد أفراد العينة وفقاً للحجم والتعداد الأصلي لكل شريحة داخل المجتمع ونسبتها إلى المجموع الكلي لمجتمع البحث .

وتعني الطبقية الشريحة أو الشرائح التي ينقسم إليها أفراد المجتمع، وتعني التناسبية أن العدد المختار من كل شريحة يجب أن يتناسب مع حجمها الفعلي ومع تمثيلها داخل المجتمع ، كما في المثال التالي :

إذا كان حجم المجتمع الأصلي هو (٢٠٠٠٠) مشاهد وكان تمثيلهم كما يلي :

(الموظفون ٤٥٠٠ + المتقاعدون ٢٥٠٠ + الطلبة ٦٠٠٠ + ربات البيوت ٣٠٠٠ + المهن الحرة ٤٠٠٠)=٢٠٠٠٠ مشاهد ، وإذا كان حجم العينة المراد اختيارها هو (٤٠٠) مشاهد، فإن تمثيلهم في العينة الطبقية التناسبية سيكون كالأتي :

{٢٠٠٠٠ حجم المجتمع ÷ ٤٠٠ حجم العينة } =٥٠ (الرقم أساس التقييم) ، وستكون العينة التناسبية كما يلي :

أ – الموظفون= (٤٥٠٠ ÷٥٠) =٩٠

ب – المتقاعدون = (٢٥٠٠÷٥٠)=٥٠

ج – الطلبة = (٦٠٠٠÷ ٥٠) = ١٢٠

د– ربات البيوت = (٣٠٠٠ ÷ ٥٠) =٦٠

هـ – مهن حرة = (٤٠٠٠ ÷٥٠) = ٨٠

٣– العينة العشوائية البسيطة : Simple Random Sample

في هذا النوع من العينات يعطي الباحث فرصة متساوية لكل فرد من أفراد المجتمع بأن يكون ضمن العينة المختارة . ويكون هذا النوع من العينات مفيد ومؤثر في حالة وجود تجانس وصفات مشتركة بين جميع أفراد المجتمع الأصلي المعني بالدراسة ، ويتم اختيار العينة العشوائية البسيطة بإحدى الطريقتين :

أ– القرعة

ب– استخدام جداول الأرقام العشوائية

ويمكن استخدام الحاسب الإلكتروني في اختيار الأرقام العشوائية بغرض سرعة الوصول إلى النماذج المطلوبة ودقة اختيارها .

٤- العينة العشوائية المنتظمة: Systematic Random Sample

يكون اختيار الوحدات في العينة المنتظمة على أساس تقسيم العدد الكلي للمجتمع على حجم العينة المطلوبة ، ومن ثم توزيع وحدات المجتمع الأصلي وبشكل متساوي ومنتظم على الرقم الناتج من ذلك التقسيم .

مثال: إذا كان العدد الكلي للمجتمع هو (٣٠٠٠) طالب وكانت العينة المطلوبة هي (١٥٠) طالب فقط فيكون توزيع الوحدات الكلية الأصلية للمجتمع كما يلي:

(٣٠٠٠÷١٥٠)- ٢٠وعلى هذا الأساس فإنه يتحدد الرقم الأول للعينة أي اسم الطالب الأول يكون أقل من الرقم (٢٠) وليكن الطالب رقم (٣) ثم يبدأ الباحث بتوزيع العينة على بقية الأسماء كما يلي : أول رقم (٣) - والرقم الثاني سيكون (٣+٢٠ -٢٣) والرقم الثالث هو (٤٣) ، ثم (٦٣)، و(٨٣) و(١٠٣) و(١٢٣)٠٠٠الخ، وهكذا حتى نصل للرقم (٢٩٨٣) أي الذي يكون تسلسله(١٥٠) .

٥- العينة العمدية أو الغرضية : Judgmental Sample

يكون الاختيار في هذا النوع من العينات على أساس حر من قبل الباحث وحسب طبيعة بحثه، بحيث يحقق هذا الاختيار هدف أو أهداف الدراسة المطلوبة . مثال ذلك اختيار الطلبة الذين تكون معدلاتهم في الامتحان النهائي جيد جدا فما فوق ، لأن هدف الدراسة هو معرفة العوامل التي تؤدي للتفوق مثلا .

٦- العينة العرضية أو عينة الصدفة: Chance Sample

و يكون الاختيار في هذا النوع من العينات سهلا إذ يعمد الباحث إلى اختيار عدد من الأفراد الذين يستطيع العثور عليهم في مكان ما وفي فترة زمنية معينة وبشكل عرضي أي عن طريق الصدفة ، كأن يذهب الباحث إلى مدرسة أو كلية أو مكتبة من المكتبات التي يتعلق البحث بها ثم يوزع الاستبيان على من يراهم موجودين أمامه .

ومن عيوب هذا النوع من العينات أنها قد لا تمثل المجتمع الأصلي تمثيلا صادقا خاصة إذا كان هناك تباين أو عدم تجانس في الخواص المطلوب دراستها في المجتمع الأصلي .

مثال: قد لا يكون القراء في اليوم الذي اختاره الباحث لزيارة المكتبة ممثلون لبقية القراء والمستفيدين الذين يستخدمون المكتبة .

التكرار frequency ، النسبة proportion ، النسبة المئوية Percentage

إذا كان لدينا ٣٠٠ طالبة في صف مؤلف من ١٢٠٠ طالب وطالبة ، فإن تكرار الطالبات هو ٣٠٠ ، ونسبة الطالبات إلى المجموع الكلي هو $\frac{٣٠٠}{١٢٠٠} = \frac{١}{٤} = ٠,٢٥$ وهذا ما تسميه (النسبة) ،

أي يكون التعبير عنها بكسر عشري ، وأخيراً إذا عبرنا عن هذه النسبة كنسبة مئوية فإننا بكل بساطة نضرب بـ ١٠٠ أي نقول ٢٥ % .

النسبة المئوية = $\frac{\text{الجزء} \times ١٠٠}{\text{الكل}}$.

التدوير rounding

التدوير : هو التقريب ، وفي الأعداد الدورية نأخذ رقمين بعد الفاصلة ، ثم ننظر للرقم الثالث إذا كان أكبر من (٥) نضيف (١) للرقم الثاني بعد الفاصلة ، وإذا كان أصغر من (٥) نبقي الرقم كما هو .

مثال : ٠,٣٣٣٣ ≅ ٠,٣٣   ٠,٦٦٦٦ ≅ ٠,٦٧

أما إذا كان الرقم الثالث هو (٥) ننظر للأرقام التي تلي الرقم (٥) فإذا كان هناك باقي ولو طفيف جداً ، فإننا نضيف (١) للرقم في المرتبة العشرية الثانية ، أما إذا كان هناك أصفار بعد الرقم (٥) فننظر للرقم في المرتبة العشرية الثانية فإذا كان رقم زوجي نبقيه على حاله ، وإذا كان رقم فردي نضيف له (١) ليصبح زوجياً .

مثال : ٠,٦٦٥٦٧٦ ≅ ٠,٦٧ ، ٦,٥٤٥٠٠١ ≅ ٦,٥٥ ، ٦,٥٤٥٠٠٠ ≅ ٦,٥٤

٦,٥٣٥٠٠٠ ≅ ٦,٥٤

مراحل البحث الإحصائي :

إن ميزة التعريف السابق لعلم الإحصاء أنه يعرف علم الإحصاء بطريقة استخدامه و بمراحله الأساسية التي لابد منها عند إجراء أي بحث إحصائي، فإذا أردنا القيام ببحث ما مستخدمين الأدوات الإحصائية فإننا و بعد التفكير العميق في المشكلة المدروسة لفهم أبعادها من أجل تحديد إطار الدراسة والبحث ، وتحديد الهدف **سنقوم بالخطوات التالية :**

١- جمع البيانات الإحصائية collecting data
٢- تصنيف وتبويب البيانات classifying data
٣- عرض البيانات showing data
٤- دراسة هذه البيانات وتحليلها باستخدام الطرق الإحصائية كمقاييس النزعة المركزية و التشتت و الارتباط وغيرها Analyzing data
٥- استخلاص النتائج Results

وتعتبر المراحل المذكورة سابقاً مرتبطة ومتعلقة ببعضها البعض إذ أن كل مرحلة تبدأ من نهاية المرحلة السابقة لها وتعتمد عليها وعلى صحة نتائجها ، لذلك فأن أي خطأ في أي مرحلة من المراحل سوف يؤثر على المراحل التي تليها وعلى صحة النتائج المستخلصة.

جمع البيانات الإحصائية :

قبل أن ندخل ببعض التفاصيل الخاصة بعملية جمع البيانات لا بد أن نتعرف أولا على أنواع البيانات الإحصائية.

أنواع البيانات الإحصائية:

تقسم البيانات الإحصائية إلى ثلاثة أنواع وهي :

١- **البيانات الكمية القابلة للقياس:** Quantitive data

وهي البيانات التي يتم الحصول عليها بنتيجة القياس بوحدة قياس معينة كالوقت (ساعة، دقيقة، يوم)، الوزن (غرام، كيلوغرام، طن)....الخ

٢- **البيانات النوعية:** Qualitative data

وهي البيانات التي تصف العناصر الإحصائية وصفاً مثل نوع التربة (رملية ، سلتية ، طينية) أو لون العيون (اسود، بني، ازرق) أو زمرة الدم (O.B,AB) أو الجنس (ذكر،أنثى) ...الخ

٣- **البيانات الترتيبية:** Ordinal data

وهي التي تقدم ترتيباً للأشياء أو العناصر تصاعدياً أو تنازلياً مثل: الأول، الثاني، الثالث...الخ، أو ترتيباً هجائياً: أ، ب،ج، أو تركيز مبيد ٠.١، ٠.٢، ٠.٣ ...الخ.

جمع البيانات الإحصائية:

يحدد موضوع الدراسة دائما نوع البيانات المطلوبة ومصدر هذه البيانات، فمثلاً إذا أردنا دراسة المتوسط الحسابي لإنتاجية محصول القمح في مدينة الحسكة. فهذا يشير لنا بضرورة جمع البيانات المتعلقة بإنتاجية المحصول، وسيكون مصدر هذه البيانات مديرية الزراعة والإصلاح الزراعي .

ويوجد هناك **ثلاث مصادر** يمكن أن تحصل منها على البيانات، وهي :

١- **المصادر التاريخية أو الرسمية:** وتتمثل عموماً في سجلات الأحوال المدنية ومكاتب الإحصاء الرسمية المتخصصة.

٢- **المصادر الداخلية:** وتتمثل في مختلف وزارات الدولة والهيئات والاتحادات والنقابات والجمعيات و الإدارات الرسمية وغير الرسمية التي تجمع عادة البيانات و المعلومات الخاصة بأمورها الداخلية .

٣- **المصادر المباشرة أو الميدانية** : ويتم اللجوء إليها في حال عدم توفر المصدر الذي يمكن أن نحصل منه على البيانات اللازمة لخصوصيتها ، فإننا سنضطر إلى جمع البيانات بأنفسنا ميدانياً .

أدوات أو وسائل جمع البيانات :

يمكن حصر أدوات أو وسائل جمع البيانات فيما يلي :

١- إجراء التجارب العلمية	٢- الاستبيان
٣- المقابلة	٤- الملاحظة

٥- المصادر والوثائق

تصنيف البيانات الإحصائية :

هناك العديد من الأسس لتصنيف البيانات الإحصائية ، ومن أهمها :

١- الأساس الجغرافي : وفيه يتم تصنيف البيانات بحسب المناطق الجغرافية المختلفة ، كأن تصنف المساحات المزروعة بالكمون حسب المناطق .

٢- الأساس النوعي : وفيه يتم تصنيف البيانات حسب النوع ، كأن تصنف النباتات بحسب معيار النضج إلى مبكرة ومتوسطة ومتأخرة النضج .

٣- الأساس الزمني ، وفيه يتم تصنيف البيانات بحسب سلسلة زمنية ، كأن تصنف تطور المساحات المزروعة من الكمون حسب السنوات .

٤- الأساس المشترك : وهو يعني اعتماد أكثر من أساس للتصنيف في الوقت نفسه ، كأن تصنف المساحات المزروعة بالكمون تبعا للسنوات والمناطق ومعيار النضج .

عرض البيانات الإحصائية :

وهناك نوعان عامان من العرض ، وهما : **العرض الجدولي والعرض البياني** حيث يجب من خلال هذا العرض شرح أكبر قدر ممكن من الظاهرة المدروسة بشكل موجز وسريع .

الفصل الثاني

إعداد : د . قصي العمر

مقاييس النزعة المركزية

Measures of central tendency

مقدمة :

تمثل مقاييس النزعة المركزية أو المتوسطات مقاييس موقع ، حيث تعطينا قيمة تمثل البيانات أكثر من باقي القيم ، وسميت كذلك لأن جميع البيانات تنزع أن تتمركز حول هذه القيمة ،وهذه المقاييس شائعة الاستخدام في كثير من مجالات الحياة ، ويهتم بدراستها المشتغلين في هذه المجالات ، فهي تهم واضعي البرامج الاقتصادية والإنتاجيةُ وذلك للتعرف على معدل الزيادة في السكان مثلاً ، أو متوسط إنتاج الدونم من المحاصيل المختلفة أو متوسط دخل الفرد ......الخ . وهي مقاييس متعددة ومتنوعة ، ومن أشهر هذه المقاييس : المتوسط أو الوسط الحسابي ، الوسيط ، والمنوال .

المتوسط الحسابي : The arithmetic mean

ويعتبر من أهم مقاييس النزعة المركزية وأكثرها استخداماً ، ويستخدم لإعطاء فكرة سريعة عن العينة والأرقام الموجودة فيها.

ويعرف المتوسط بأنه : **مجموع القيم مقسوماً على عددها .**

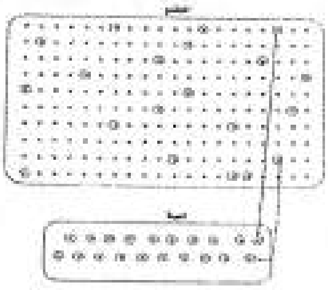
ويمكن التعبير عن هذا التعريف رياضياً بالعلاقة التالية :

$$\overline{x} = \frac{x_1 + x2 ....+ x_n}{n}$$

أو

$$\overline{x} = \sum_{i=1}^{n} x_i \ /n$$

$\sum$: مجموع



المتوسط الحسابي

[illegible]

مثال : احسب متوسط البيانات التالية:

3 , 5 , 1 , 4 , 2 , 3 , 5 , 1 , 4 , 2 , 3 , 5 , 1 , 4 , 2, 3 , 5 , 1 , 4 , 2

$$\overline{x} = \frac{3+5+1+4+2+3+5+1+4+2+3+5+1+4+2+3+5+1+4+2}{20}$$

$$\overline{x} = \frac{60}{20} = 3$$

ومن مزايا المتوسط ما يلي :

١ -سهل الحساب وسهل التعريف .

٢ -وحيد لمجموعة من البيانات .

٣ -يدخل في الاعتبار عند حسابه جميع البيانات .

٤ -مجموع انحرافات القيم عن متوسطها الحسابي يساوي الصفر .

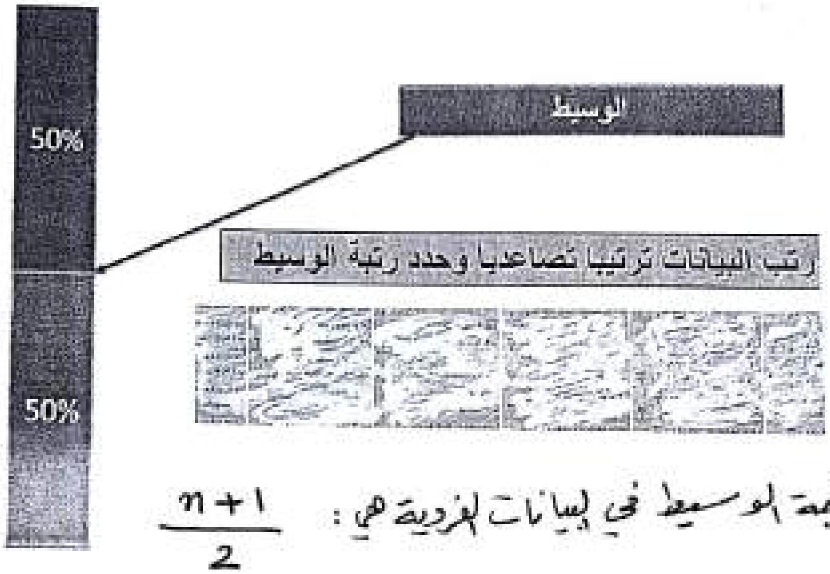
ومن عيوب المتوسط الحسابي ما يلي :

١ -تأثره الشديد بالقيم الشاذة أو المتطرفة .

٢ -لا يمكن حسابه لبيانات نوعية مثل نوع الترب ، لون العيون ، زمر الدم......الخ .

الوسيط : The Median

يعتبر الوسيط أحد مقاييس النزعة المركزية وهو يعتبر متوسط مكاني حيث يكون عدد الأفراد التي قيمتها أقل منه مساوي لتلك التي قيمتها أكبر منه ، ويعرف الوسيط لمجموعة من الأعداد المرتبة ترتيبا تصاعديا أو تنازليا بأنه العدد الأوسط منها إذا كان عددها فرديا، وهو الوسط الحسابي للعددين الأوسطين إذا كان عددها زوجيا.



رتبة الوسيط في البيانات الفردية هي : $\frac{n+1}{2}$

رتبة الوسيط في البيانات الزوجية : $\frac{n}{2}$ للقيمة الأولى ، $\frac{n}{2}+1$ للقيمة الثانية

الوسيط

مثال (1): أوجد قيمة الوسيط للبيانات التالية:

$27 , 12 , 15 , 300 , 22 , 18 , 14 , 24 , 17$

الحل: 1- ترتيب البيانات: $12 , 14 , 15 , 17 , \textcircled{18} , 22 , 24 , 27 , 300$

الوسيط

2- تحدد موقع الوسيط (رتبته) كالتالي:

بما أن عدد البيانات : $n = 9$ وهو عدد فردي، إذن رتبة الوسيط هي:

$$\frac{n+1}{2} \text{ أي } \frac{9+1}{2} = 5$$

3- إذاً قيمة الوسيط هي القيمة الخامسة في البيانات المرتبة وهي:

$$Med = 18$$

الوسيط

مثال (2): أوجد قيمة الوسيط للبيانات التالية:

$3 , 9 , 12 , 16 , 14 , 33 , 10 , 15 , 20 , 7$

الحل: 1- ترتيب البيانات: $3 , 7 , 9 , 10 , \textcircled{12} , \textcircled{14} , 15 , 16 , 20 , 33$

القيمتين الوسطيتين

2- تحديد موقع الوسيط: عدد البيانات $n = 10$ ، بما أن عدد البيانات عدد زوجي، فإن

الوسيط له رتبتين: $\frac{10}{2} = 5$ و $\frac{10}{2} + 1 = 6$

3- تحدد قيمة الوسيط، وهي متوسط القيمتين الوسيطتين: الخامسة والسادسة، وهي:

$$\frac{12+14}{2} = 13$$

المنوال : The Mode

وهو القيمة التي تتكرر أكثر من غيرها في المفردات الإحصائية ، أو القيمة الأكثر تكرارا أو شيوعا ولهذا يطلق عليه أحيانا الشائع أو القيمة الشائعة ، ويعتبر المنوال من أسهل مقاييس النزعة المركزية لأن الحصول عليه يتم بالنظر أكثر مما يتم بالحساب .

وقد يكون للمجموعة منوال واحد أو أكثر من منوال أو قد لا يوجد لها منوال.

مثال (١) :

أوجد المنوال للأعداد التالية: ٢٥، ٢٠، ٢٨، ٢٠، ٢٠، ١٥، ٣٠.

الحل:

بما أن القيمة (٢٠) تتكرر أكثر من غيرها و بناء على تعريف المنوال إذاً يكون المنوال لهذه المجموعة يساوي (٢٠).

مثال (٢) :

ما المنوال لمجموعة الأعداد التالية: ٥٦، ٤٠، ٥٦، ٥٦، ٧٢، ٣٤، ٤٩، ٧٢، ٧٢؟

الحل:

يوجد لهذه المجموعة منوالان هما: ٥٦، ٧٢ لأن كلا منهما يتكرر بنفس عدد المرات التي يتكرر فيها الآخر.

مثال (٣) :

ما المنوال لمجموعة الأعداد: ٢١، ١٦، ٢٤، ٩، ٨٤، ٣٨؟

الحل: لا يوجد منوال لهذه المجموعة لأن كلا منها يتكرر مرة واحدة

الفصل الثالث

إعداد : د . قصي العمر

مقاييس التشتت أو الاختلاف

Measures of Dispersion

مقدمة :

إن مقاييس النزعة المركزية غير كافية لتحديد صفات البيانات الإحصائية ، فربما يكون لدينا ظاهرتان متساويتان في الوسط الحسابي والوسيط إلا أنهما مختلفتان.

فمثلاً إذا كانت درجات الحرارة في بلد ما خلال الليل و النهار هي : ٢٢، ٢٨, ٢١ ، ١٩، ١٠ فيكون معدل أو متوسط درجات الحرارة ٢٠ .

وإذا كانت درجات الحرارة في بلد آخر هي: ٢٤، ٢٩، ٢١، ١٦، ١٠ فيكون معدل أو متوسط درجات الحرارة فيها ٢٠ أيضاً .

وهذا يعني أن معدل درجات الحرارة في البلدين متساوي ، إلا أنه بالنظر إلى مفردات البيانات في كل من البلدين نجد اختلافاً بينهما ، وهذا يعني أن الوسط الحسابي لا يكفي لوصف البيانات أو للحكم على تشابهها.

ويقاس التباعد بدرجة انتشار البيانات حول معدل أو وسط ، وكلما كان التغير صغيراً اعتبرت البيانات متجانسة ، وكلما كبر التغير اعتبرت البيانات متباعدة أو متغايرة. وعلى ما تقدم فإننا بحاجة إلى مقاييس توضح مقدار تشتت أو تبعثر أو اختلاف البيانات عن بعضها البعض لاكتمال صورة وصف لبيانات ، وتقوم مقاييس التشتت بهذا الدور .

ومن أهم هذه المقاييس ما يلي : ١- المدى المطلق ٢- التباين ٣- الانحراف المعياري ٤- الخطأ القياسي ٥- معامل الاختلاف

وفيما يلي شرح مبسط عن كل واحد من هذه المقاييس .

المدى المطلق : The range

يعرف المدى للبيانات بأنه الفرق بين أكبر قيمة وأصغر قيمة.

$$R = X_{max} - X_{min}$$

X_{max} : أكبر قيمة ، X_{min} : أصغر قيمة

ويستخدم المدى عادةً لإعطاء فكرة سريعة أولية عن طبيعة توزيع المفردات الإحصائية لأنه يمتاز ببساطته وسهولة حسابه، لكنه لا يعكس أثر جميع المشاهدات لأنه يعتمد على أكبر وأصغر قيمتين فقط

المدى

الفرق بين أكبر قيمة وأصغر قيمة في مجموعة ما

A : 5 , 6 , 8 , 10 , 12 , 14 , 15

Range = 15-5 =10

B : 1 , 2 , 5 , 10 , 15 , 18 , 19

Range = 19 -1 =18

التباين والانحراف المعياري Variance and Standard Deviation

ويوضح هذان المقياسان مدى تقارب أو تباعد الدرجات عن المتوسط الحسابي ، و هما من أدق مقاييس التشتت وأكثرها استعمالا ، ويعرف التباين بأنه : متوسط مربعات انحرافات القيم عن متوسطها و يرمز له بالرمز S^2 .بينما يعرف الانحراف المعياري بأنه : الجذر التربيعي الموجب لمتوسط مربعات انحرافات القيم عن متوسطها و يرمز له بالرمز S ، ويمكن التعبير عن التعريفين رياضياً بالصيغتين التاليتين :

التباين داخل متغير x $S^2 = \frac{\Sigma(x-\overline{x})^2}{n-1}$ الانحراف المعياري $S = \sqrt{\frac{\Sigma(x-\overline{x})^2}{n-1}}$

التباين σ^2 S^2

مثال : أوجد تباين العينة الممثلة بالبيانات التالية:

5 8 4 7 4 2

التباين

الحل :

$$\bar{X} = \sum_{i=1}^{n} x_i / n = \frac{30}{6} = 5$$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

$$s^2 = \frac{(5-5)^2 + (5-8)^2 + (5-4)^2 + (5-7)^2 + (5-4)^2 + (5-2)^2}{6-1}$$

$$s^2 = \frac{(0)^2 + (-3)^2 + (1)^2 + (-2)^2 + (1)^2 + (3)^2}{5}$$

$$s^2 = \frac{0+9+1+4+1+9}{5} = \frac{24}{5} = 4.8$$

$$S = \sqrt{S^2} = \sqrt{4.8} = 2.19 \longrightarrow$$ هو الانحراف المعياري

الخطأ القياسي : Standard Error

يعبر الخطأ القياسي عن مقياس يمكن التعرف بواسطته عن مدى تشتت المتوسطات لعدد من العينات المأخوذة من نفس المجتمع ، فإذا كانت قيمته كبيرة دل ذلك على تبعثر المتوسطات ، وإذا كانت قيمته صغيرة اعتبر ذلك دليلاً على تركيز المتوسطات حول متوسط المجتمع . وقد اعتاد الاحصائيون لاستخراج قيمة الخطأ القياسي بحسابه من العينة التي يعمل في إطارها معتمداً في ذلك على قيمة الانحراف القياسي للعينة وفق المعادلة التالية :

$$S\bar{X} = S / \sqrt{n}$$

$S\bar{X}$: الخطأ القياسي .

S : الانحراف القياسي أو المعياري للعينة .

n : عدد أفراد العينة

معامل الاختلاف : Coefficient of Variability

وهو أحد مقاييس التشتت النسبية ، ويستخدم لمقارنة المجتمعات أو العينات المختلفة ، وذلك عن طريق التعبير عن الانحراف القياسي للعينة أو المجتمع في صورة نسبة مئوية من المتوسط الحسابي ، وهو مقياس مجرد من وحدة قياس البيانات ، إذ يمكن مقارنة تشتت أطوال النباتات في سلالة الشعير مقيسة بوحدات من السنتمترات بتشتت أوزان محصول من نفس السلالة مقدراً بالكيلوغرام .

ويمكن التعبير رياضياً عن هذا المقياس وفق المعادلة التالية :

$$C.V = S / \overline{X} * 100$$

C.v : معامل الاختلاف

S : الانحراف المعياري

X : المتوسط الحسابي

* : ضرب

/ : مقسوماً على

مثال :

بلغ متوسط أطوال عينة من نباتات الشعير 90cm وانحراف معياري 9 ، كما بلغ متوسط أوزان عينة من نباتات الشعير 30 kilogram وانحراف معياري 2 ، والمطلوب : أي العينتين أكثر اختلافاً أو تغيراً ؟

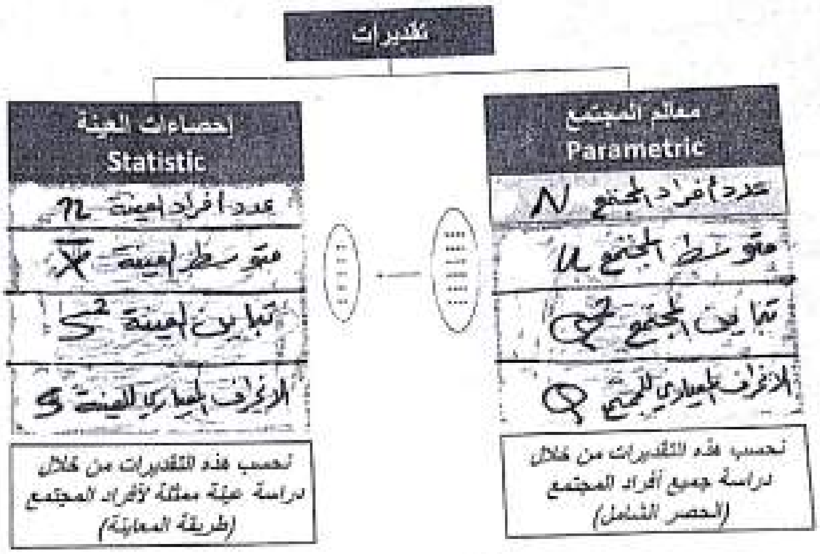
الحل :

نحسب معامل الاختلاف للعينة الأولى (الأطوال) ومعامل الاختلاف للعينة الثانية (الأوزان) ، بحيث يكون المعامل الأكبر هو الأكثر اختلافاً .

C.V 1 = 9 / 90 * 100 = 10 %

C.V 2 = 2 / 30 * 100 = 6.67 %

أي أن العينة الأولى أكثر اختلافا من العينة الثانية .



الفصل الرابع

إعداد : أ.د. قصي ...
م. علي ملي

الارتباط والانحدار
Correlation & Regrresion

مقدمة :

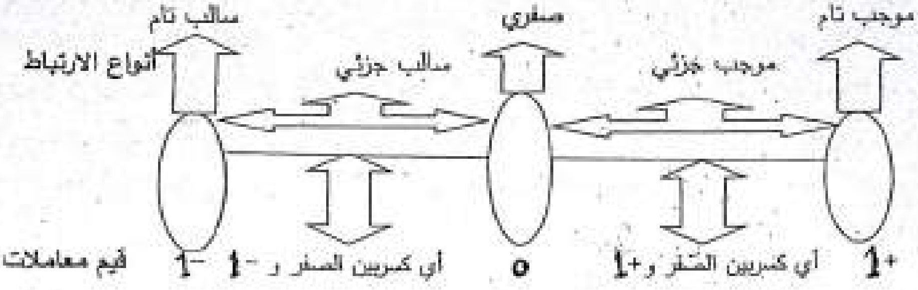
إن اهتمامنا في المحاضرات السابقة كان موجهاً نحو الطرق العديدة التي يمكن بها وصف البيانات لمتغير واحد (... أطوال بيانات ، علامات مجموعة من الطلبة في أحد الامتحانات – أجور عمال....الخ) مثل تنظيم البيانات في جداول توزيع تكراري ورسوم بيانية ، بالإضافة إلى استخراج مقاييس النزعة المركزية ومقاييس التشتت ، وفي هذه الحالات تكون دراستنا عن متغير واحد . ولكن هناك العديد من المسائل في العلوم التطبيقية تتطلب أن تذهب إلى ما هو أبعد من وصف متغير واحد ، ألا وهو تحديد العلاقات بين متغيرين أو أكثر ، فقد نرغب مثلاً بدراسة العلاقة بين سعر السلعة والكمية المطلوبة منها ، أو معرفة العلاقة بين علامة الطالب في الامتحان النهائي وعدد الساعات التي قضاها الطالب في التحضير للامتحان ، أو معرفة العلاقة بين الإنفاق والدخل العائلي ، أو معرفة العلاقة بين عدد مرات ممارسة نوع معين من الرياضة البدنية ومستوى الكلسترول في الدم ، أو معرفة العلاقة بين كميات السماد المستخدمة وكمية الإنتاج من محصول معين ، مثل هذه الأسئلة وإجاباتها تندرج تحت موضوعين من أهم مواضيع الإحصاء وهما الارتباط والانحدار Correlation & Regression

فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط ، وإذا كان اهتمامه بدراسة أثر أحد المتغيرين على الآخر استخدم لذلك أسلوب تحليل الانحدار .

الارتباط :

تدل مقاييس الارتباط على العلاقة بين متغيرين أو أكثر أو ميل ظاهرتين إلى التغير معاً ، وهذا التغير يكون إما في اتجاه واحد ويسمى اتجاه طردي أو ارتباط موجب ، أو في اتجاهين مختلفين ويطلق عليه اتجاه عكسي أو ارتباط سالب ، وسواء أكانت العلاقة بين المتغيرين طردية أو عكسية فدرجة هذه العلاقة تسمى معامل الارتباط **Coefficient of Correlation** ، ويرمز له بالرمز "r" ، وحدود هذا المعامل بين +1 و -1 وهذا يعني أن : $+1 \geqslant r \geqslant -1$

وهناك خمسة أنواع للارتباط يترتب عليها خمس قيم لمعاملات الارتباط ، كما هو موضح بالشكل التالي :



فمعامل الارتباط التام يعني أن زيادة أو نقصان أحد المتغيرين يتبعها زيادة أو نقصان الآخر ولكن بنفس الدرجة أو النسبة . والجزئي يعني أن زيادة أو نقصان أحد المتغيرين يتبعها زيادة أو نقصان الآخر ولكن ليس بنفس الدرجة أو النسبة . أما الصفري فهو يعني انعدام العلاقة بين المتغيرين ، أي أنه لا توجد علاقة أو ارتباط بين المتغيرين .

وفي حالة الارتباط الموجب تكون العلاقة بين المتغيرين علاقة طردية بمعنى أن الزيادة في أحدهما يتبعها زيادة في الآخر مثل : الإنتاجية وكمية الحوافز ، الوزن وضغط الدم ، العمر والنضج ، الإنفاق الاستهلاكي والدخل " ؛ وفي حالة الارتباط السالب تكون العلاقة بين المتغيرين علاقة عكسية بمعنى أن الزيادة في أحدهما يتبعها نقصان في الآخر مثل : الإنتاجية والإصابة [illegible] ، الرضا الوظيفي والغياب عن العمل ، التدين والتوجه نحو الحياة العصرية ، الإنتاج والتكلفة ، وقد لا توجد علاقة إطلاقاً وفي هذه الحالة يكون معامل الارتباط متساوياً الصفر مثل العلاقة بين طول الفرد ومستوى ثقافته أو العلاقة بين الوزن ومستوى التحصيل الدراسي .

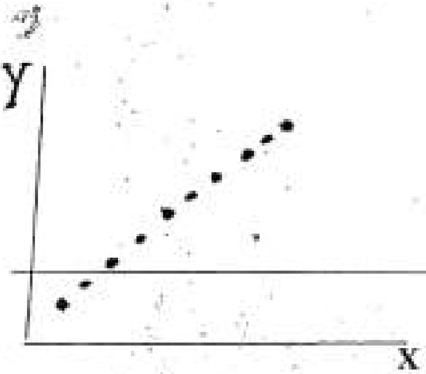
لوحة الانتشار Scatter diagram

إن إحدى الطرق التي تساعدنا على معرفة وجود علاقة بين متغيرين x و y هي لوحة الانتشار ، حيث نرسم إحداثياً أفقياً (إحداثي x) وإحداثياً عمودياً عليه (احداثي y) ونرصد أزواج المشاهدات المرتبة :

(x1, y1), (x2, y2) ، و ....(xn , yn) المعطاة لدينا على المستوى xy فنحصل على لوحة الانتشار

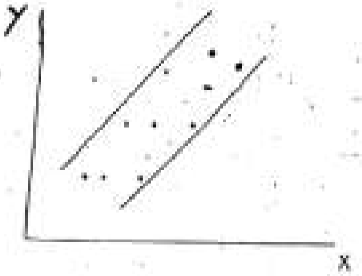
وتبين هذه اللوحة بشكل جيد ما إذا كان هناك علاقة بين المتغيرين أو عدم وجودها ، كذلك تبين لوحة الانتشار كون العلاقة بين المتغيرين طردية أو عكسية ، ومقدار هذه العلاقة بشكل تقريبي .

والأشكال التالية توضح لوحات الانتشار لأنواع الارتباط الخمسة المذكورة سابقاً .



ارتباط تام موجب

Positive & complete correlation



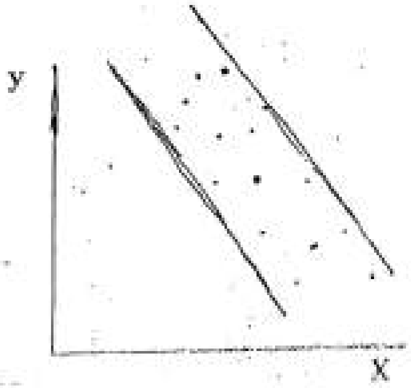
ارتباط جزئي موجب

Positive & Partial correlation

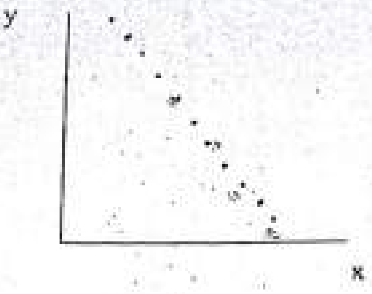
عدم وجود علاقة أو ارتباط صفري

Zero correlation

ارتباط سالب جزئي



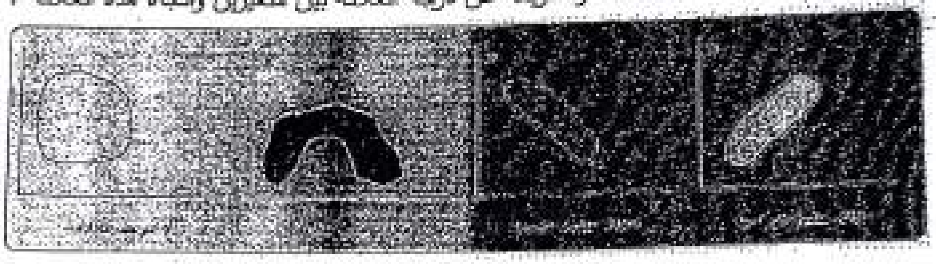
Negative & Partial correlation



ارتباط سالب تام

Negative & complete correlation

وبالنظر إلى هذه الأشكال نأخذ فكرة سريعة عن درجة العلاقة بين متغيرين واتجاه هذه العلاقة .

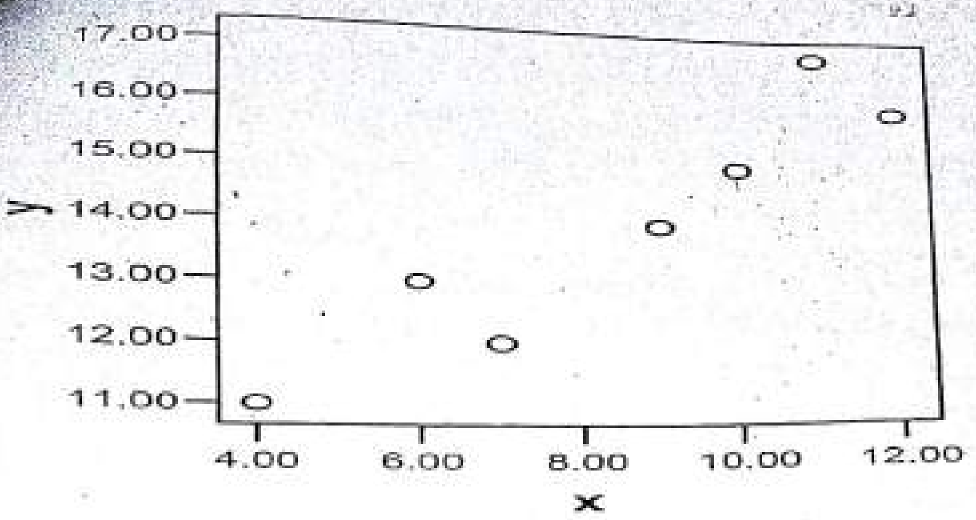


مثال : سجل أحد المحاضرين عدد الساعات التي قضاها 7 طلاب تم اختيارهم عشوائياً من طلاب كلية الزراعة في التحضير لامتحان الإحصاء ، وعلاماتهم في ذلك الامتحان في أول الأسبوع التالي ، فوجدها كالتالي :

عدد الساعات x	6	12	4	11	10	7	9
العلامة y	13	16	11	17	15	12	14

المطلوب : ارسم لوحة الانتشار ، ووضح العلاقة بين العلامة وعدد ساعات الدراسة ؟

الحل : افرض x عدد الساعات ، y العلامة ، ثم ارصد النقط التي إحداثياتها الأزواج المرتبة (6 و 13) ، (12 ، 16) .... (9 ، 14) :



يظهر من لوحة الانتشار أنه يمكن اعتبار العلاقة بين المتغيرين طردية أو موجبة جزئية ، أي كلما زادت x زادت y والعكس صحيح ، كذلك فإنها قريبة من الواحد الصحيح .

معاملات الارتباط :

هناك ثلاثة أشكال من معاملات الارتباط ، وهي :

١- الارتباط البسيط Simple Correlation : وهو الارتباط بين ظاهرتين فقط .

٢- الارتباط المتعدد Multiple Correlation : وهو الارتباط بين ظاهرة (متغير) وعدد آخر من الظواهر في وقت واحد.

٣- الارتباط الجزئي Partial Correlation: وهو الارتباط بين ظاهرة وظاهرة أخرى مع وجود ظواهر أخرى يتم تثبيتها عند الدراسة.

وسوف نقتصر دراستنا على الشكل الأول ، وهو الارتباط البسيط ، حيث يوفر لنا علم الإحصاء عدداً من الأساليب لحساب معامل الارتباط بين المتغيرين ، ويقوم كل أسلوب على خصائص المتغيرات التي نتعامل معها . ومن أشهر هذه الأساليب معامل ارتباط بيرسون

معامل ارتباط بيرسون :

وهو أكثر أنواع معاملات الارتباط استخداماً و أكثرها شيوعاً في البحوث ، وهو نسبة إلى العالم كارل بيرسون ويستخدم في البيانات الكمية .

وهناك عدة طرق يمكن استخدامها لحساب معامل ارتباط بيرسون تتمثل في ما يلي :

١- طريقة الانحرافات .

٢- طريقة البيانات الأصلية أو الدرجات الخام .

٣- طريقة التغاير .

٤- طريقة الدرجات المعيارية .

وعلى العموم تعد **طريقة الانحرافات وطريقة البيانات الأصلية** هي الأكثر شيوعاً واستخداماً في حساب معامل ارتباط بيرسون لأنها تتعامل مع المشاهدات الأصلية ولا تتطلب أي نوع من التغيير أو التحويل ، وسوف يتم التعرض لطريقة البيانات الأصلية فقط لسهولتها

*** طريقة البيانات الأصلية :**

يمكن حساب معامل ارتباط بيرسون وفق الصيغة الرياضية التالية :

$$\boxed{r=\frac{\sum xy-\frac{\sum x\sum y}{n}}{\sqrt{\left(\sum x^2-\frac{(\sum x)^2}{n}\right)\left(\sum y^2-\frac{(\sum y)^2}{n}\right)}}}$$

مثال : احسب معامل إرتباط بيرسون بين إيداعات خمسة من العملاء في المصرف المركزي (X) والمصرف التجاري (Y) ؟

	x	y	xy	x^2	y^2
	4	2	8	16	4
	7	3	21	49	9
	8	6	48	64	36
	10	7	70	100	49
	15	10	150	225	100
$\sum$	44	28	297	454	198

الحل :

1- نحسب مجموع قيم (x) وقيم (y) كما هو مبين في العمودين (1) و (2)

2- نوجد مربع قيم (x) ومربع قيم (y) كما هو مبين في العمودين (4) و (5)

3- نحسب ناتج ضرب قيم (x) بقيم (y) كما هو مبين بالعمود الثالث .

4- n تمثل عدد أزواج البيانات وهي في مثالنا 5 .

5- نعوض في الصيغة الرياضية السابقة كالتالي :

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}}$$

$$= \frac{297 - \frac{(44)(28)}{5}}{\sqrt{\left(454 - \frac{(44)^2}{5}\right)\left(198 - \frac{(28)^2}{5}\right)}}$$

$$= \frac{297 - 246.4}{\sqrt{66.8 * 41.2}} = \frac{50.6}{52.46} = 0.96$$

إذا كانت [illegible] نقاط معين، فإن [illegible] متغيرين (أي أن العلاقة بينهما ضعيفة جداً)

كما يمكن حساب معامل الارتباط بطريقة الانحرافات وفق المعادلة التالية

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 * \sum_{i=1}^{n}(y_i-\bar{y})^2\right.}}$$

★ ضرب

[illegible] بين المتغيرين

- <u>اختبار معنوية معامل الارتباط (الأهمية الإحصائية والدلالة الإحصائية لمعامل الارتباط):</u>

❖ <u>النظرية الفرضية</u> التي توضع تحت الاختبار هي فرضية العدم (H0: لا توجد علاقة معنوية بين الصفتين الكميتين).

$H0: \rho = 0$ ⟵ فرضية العدم Null-hypothesis

$H1: \rho \neq 0$ ⟵ الفرضية البديلة Alternative hypothesis

P : لفظ لاتيني يعرف برو وهو معامل ارتباط المجتمع الاصلي.

- إذا كانت r المحسوبة > r الجدولية عند مستوى المعنوية ودرجات الحرية المحددة (n-2)

⟸ توجد علاقة معنوية بين الصفتين، أي إن الارتباط بين الصفتين يختلف معنوياً عن الصفر ولا يعود الاختلاف بين الصفتين الكميتين للصدفة، ومعامل الارتباط له أهمية إحصائية، والعينة تمثل مجتمعها، ومعامل ارتباط المجتمع الذي سحبت منه العينة لا يساوي الصفر (ترفض فرضية العدم وتقبل الفرضية البديلة).

- إذا كانت r المحسوبة < r الجدولية عند مستوى المعنوية ودرجات الحرية المحددة (n-2)

⟸ لا توجد علاقة معنوية بين الصفتين، أي إن الارتباط بين الصفتين لا يختلف معنوياً عن الصفر، ومعامل الارتباط ليس له أهمية إحصائية، والعينة لا تمثل مجتمعها، ومعامل ارتباط المجتمع الذي سحبت منه العينة تساوي الصفر (تقبل فرضية العدم وترفض الفرضية البديلة).

<u>ملاحظة هامة:</u> لجعل معامل الارتباط معنوي يجب أن نسعى أن تكون n كبيرة بقدر الإمكان

- **معامل التحديد Coefficient of Determination:**

يسمى مربع معامل الارتباط بمعامل التحديد ويرمز له (R square)، ويفيد هذا المعامل في تحديد النسبة المئوية لتأثير العامل المستقل على العامل التابع(مقدار التباين في أحد المتغيرات التي يمكن تفسيره من خلال المتغير الأخر)، فلو فرضنا إننا أوجدنا قيمة معامل الارتباط بين كمية السماد وبين كمية الإنتاج من محصول معين وكانت(0.4)، ففي هذه الحالة يكون معامل التحديد (0.16)، ويعني ذلك أن التغيرات التي تحدث في الإنتاج تكون ناتجة عن تأثير السماد بنسبة (16 %) فقط أما بقية النسبة (84 %) ترجع لعوامل أخرى لم تدخل في التحليل ولم تكن متحكم بها، قد تكون كمية المياه المعطاة أو حرارة الجو أو ميعاد الزراعة ....الخ التي تؤثر بدورها على التغيرات التي تطرأ على الإنتاج.

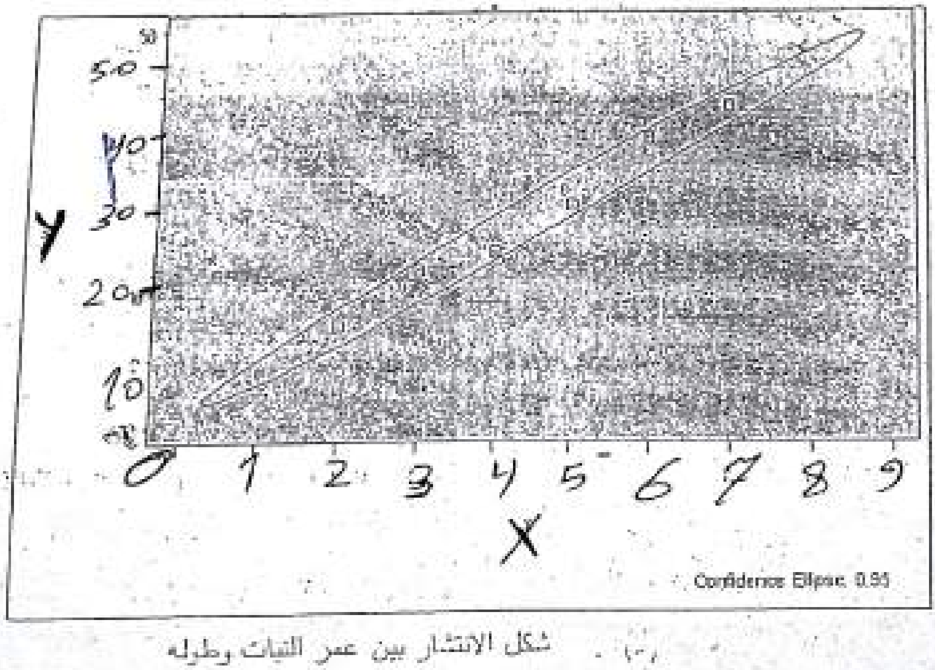
كما يفيد معامل التحديد في معرفة مدى قوة أو ضعف الارتباط ، فإذا كانت قيمته أكبر من 0.75 فالارتباط قوي جداً وإذا كانت قيمته بين 0.75 - 0.50 فالارتباط قوي ، أما إذا كانت قيمته أصغر من 0.50 فالارتباط ضعيف .

مثال : في تجربة لدراسة العلاقة بين عمر النبات بالأسبوع وطوله بـ سم تم الحصول على النتائج التالية:

عمر النبات بالأسبوع x	1	2	3	4	5	6	7
طول النبات بـ سم y	5	13	16	23	33	38	40

المطلوب:

- احسب معامل الارتباط بين عمر النبات وطوله وفسره.
- اختبر معنوية معامل الارتباط وفسره.
- احسب قيمة معامل التحديد وفسره.



(٣) شكل الانتشار بين عمر النبات وطوله

الحل:

x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
1	5	-3	-19	9	361	57
2	13	-2	-11	4	121	22
3	16	-1	-8	1	64	8
4	23	0	-1	0	1	0
5	33	1	9	1	81	9
6	38	2	14	4	196	28
7	40	3	16	9	256	48
$\sum x_i=28$	$\sum y_i=168$	0	0	28	1080	172

$$\bar{x}=\frac{\sum x_i}{n}=\frac{28}{7}=4\,week$$

$$\bar{y}=\frac{\sum y_i}{n}=\frac{168}{7}=24\,cm$$

$$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}\rightarrow r=\frac{172}{\sqrt{28\times1080}}=0.98$$

1- الارتباط طردي (موجب) وقوي جداً (تغير العمر يؤدي وبشكل قوي إلى تغير الطول).

2- اختبار معنوية معامل الارتباط:

- **طريقة استخراج r الجدولية:**

- يتم إنشاء خط شاقولي من مستوى المعنوية المطلوب.
- يتم إنشاء خط أفقي من درجات الحرية (n-2).
- يتقاطع الخطان عند رقم يمثل قيمة r الجدولية.
- يلاحظ أن قيمة r الجدولية عند مستوى 0.05 ودرجة حرية(7-2=5) = 0.754.
- يلاحظ أن قيمة r الجدولية عند مستوى 0.01 ودرجة حرية(7-2=5) = 0.874.

يُلاحظ أنّ r المحسوبة $>$ r الجدولية عند مستوى المعنوية 5% و 1% ودرجات الحرية 5 ⬅ توجد علاقة معنوية جداً بين الصفتين، أي أنّ الارتباط بين الصفتين يختلف معنوياً عن الصفر ولا يعود الاختلاف للصدفة، ومعامل الارتباط له أهمية إحصائية، والعينة تمثل مجتمعها، ومعامل ارتباط المجتمع الذي سحبت منه العينة لا يساوي الصفر (ترفض فرضية العدم وتقبل الفرضية البديلة).

$$R\,square = r^2 = (0.98)^2 = 0.96$$

- ‎96% من الزيادة في طول النبات بـ سم ناتج عن زيادة عمر النبات بالأسبوع، و 4% يعود إلى ظروف أخرى غير متحكم بها (96% من التباين في الطول يُفسر من خلال العمر).
- الارتباط قوي جداً لأنه أكبر من 0.75 .

قيم معامل الارتباط عند مستويات معنوية مختلفة

α n	.1	.05	.02	.01	.001
1	.98769	.99692	.999507	.999877	.9999988
2	.90000	.95000	.98000	.990000	.99900
3	.8054	.8783	.93433	.95873	.99116
4	.7293	.8114	.8822	.91720	.97406
5	.6694	.7545	.8329	.8745	.95074
6	.6215	.7067	.7887	.8343	.92493
7	.5822	.6664	.7498	.7977	.8982
8	.5494	.6319	.7155	.7646	.8721
9	.5214	.6021	.6851	.7348	.8471
10	.4973	.5760	.6581	.7079	.8233
11	.4762	.5529	.6339	.6835	.8010
12	.4575	.5324	.6120	.6614	.7800
13	.4409	.5139	.5923	.6411	.7603
14	.4259	.4973	.5742	.6226	.7420
15	.4124	.4821	.5577	.6055	.7246
16	.4000	.4683	.5425	.5897	.7084
17	.3887	.4555	.5285	.5751	.6932
18	.3783	.4438	.5155	.5614	.6787
19	.3687	.4329	.5034	.5487	.6652
20	.3598	.4227	.4921	.5368	.6524
25	.3233	.3809	.4451	.4869	.5974
30	.2960	.3494	.4093	.4487	.5541
35	.2746	.3246	.3810	.4182	.5189
40	.2573	.3044	.3578	.3932	.4896
45	.2428	.2875	.3384	.3721	.4648
50	.2306	.2732	.3218	.3541	.4433
60	.2108	.2500	.2948	.3248	.4078
70	.1954	.2319	.2737	.3017	.3799
80	.1829	.2172	.2565	.2830	.3568
90	.1726	.2050	.2422	.2673	.3375
100	.1638	.1946	.2301	.2540	.3211

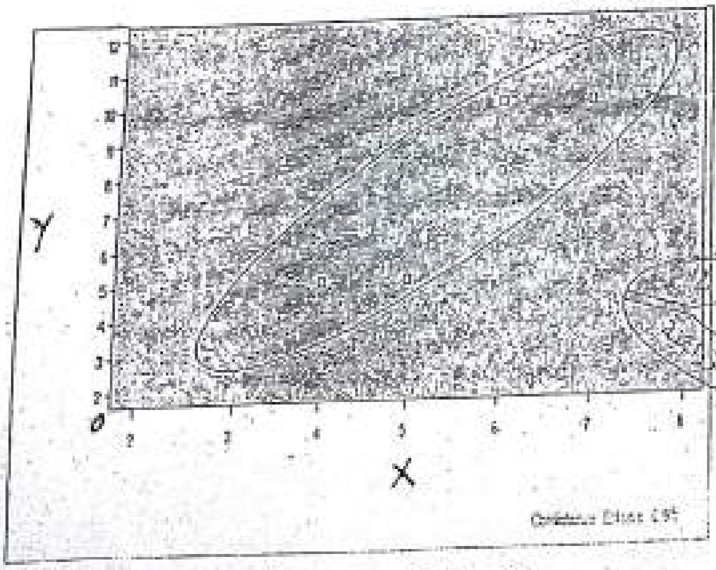
مثال :

- وجد أن هناك علاقة بين عدد الجوزات في نبات محصول القطن وعدد الأفرع الثمرية لنفس النبات:

عدد الأفرع x	4	5	6	7	6	4	4	6
عدد الجوزات y	6	5	8	10	8	6	5	10

المطلوب:

- حساب مدى قوة هذه العلاقة (معامل الارتباط).
- اختبر معنوية معامل الارتباط.
- احسب معامل التحديد وفسر معناه.



شكل الانتشار بين عدد الأفرع الثمرية وعدد الجوزات

الحل:

x_i	y_i	x_i^2	y_i^2	x_iy_i
4	6	16	36	24
5	5	25	25	25
6	8	36	64	48
7	10	49	100	70
6	8	36	64	48
4	6	16	36	24
4	5	16	25	20
6	10	36	100	60
$\sum x_i = 42$	$\sum y_i = 58$	230	450	319

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right]\left[\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right]}} = \frac{319 - \frac{42 \times 58}{8}}{\sqrt{\left[230 - \frac{(42)^2}{8}\right]\left[450 - \frac{(58)^2}{8}\right]}} = 0.86$$

أي أن الارتباط موجب (طردي) وقوي.

- يلاحظ أن قيمة r الجدولية عند مستوى 0.05 ودرجة حرية(6=2-8) = 0.7067.
- يلاحظ أن قيمة r الجدولية عند مستوى 0.01 ودرجة حرية(6=2-8) = 0.8343.
- يُلاحظ أن r المحسوبة > r الجدولية عند مستوى المعنوية 5% و 1% ودرجات الحرية 6

⬅ توجد علاقة معنوية جداً بين الصفتين، أي إنَ الارتباط بين الصفتين يختلف معنوياً عن الصفر وأنّ هذه الاختلافات ليست راجعة للصدفة.

مثال :

في تجربة لدراسة العلاقة بين كمية الإنتاج لنبات ما والسماد العضوي المضاف، تم الحصول على النتائج التالية:

كمية السماد x طن/ هـ	14	7	13	8	10	5	6	9
الإنتاج y طن/ هـ	11	6	9	6	8	4	5	7

والمطلوب:

- حساب مدى قوة هذه العلاقة (معامل الارتباط)، و حساب معامل التحديد وتفسير معناه

الجواب: $r = 0.982$ الارتباط موجب وقوي جداً.

$$R\,square = r^2 = (0.982)^2 = 0.96$$

- 96% من الزيادة في كمية الإنتاج ناتج عن الإضافات السمادية و 4% تعود إلى ظروف أخرى غير السماد كالري، والظروف البيئية والصنف (96% من التباين في كمية الإنتاج يفسر من قبل السماد).

الإنحدار
(Regression Analysis)

- مفهوم الانحدار:

إنّ نموذج الانحدار يعبّر عن العلاقة بين متغيرين، أحدهما يعبر عنه بالمتغيّر التابع (Dependent Variable)، والآخر هو واحد أو أكثر من المتغيرات المستقلة (Independent Variable)، وما يهمنا هو معرفة مدى تأثير المتغير المستقل (X) على المتغير التابع (Y)، وحساب مقدار التغير في المتغير التابع عندما يتغير المتغير المستقل بمقدار معين زيادة أو نقصاناً، وتقدير قيمة المتغير التابع عند قيمة محددة للمتغير المستقل (التنبؤ بقيمة المتغير التابع)، فالانحدار هو عبارة عن تمثيل العلاقة بين متغيرين بمعادلة رياضية وهذه المعادلة الرياضية تأخذ أشكالاً مختلفة، فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط أما إذا كان اهتمامه بدراسة تأثير أحد المتغيرين على الآخر استخدم لذلك أسلوب تحليل الانحدار.

- الانحدار الخطي البسيط Simple Linear Regression:

إذا احتوى نموذج الانحدار على متغير مستقل واحد، يعرف عندئذ بالانحدار البسيط Simple Regression Model، وإذا احتوى أكثر من متغير مستقل، فيعرف بالانحدار المتعدد Multiple Regression Model، كما إن النموذج قد يكون خطياً Linear Model، وقد يكون غير خطي Non Linear Model.

أمثلة على الانحدار البسيط:

1- الزيادة في كمية الإنتاج لمحصول ما (Y) تكون ناتجة عن زيادة عدد النباتات في وحدة المساحة (X). [دراسة تأثير زيادة عدد النباتات على كمية الإنتاج].

2- انخفاض كمية الإنتاج لمحصول ما (Y) تكون ناتجة عن كثرة الإصابة بالأمراض أو الحشرات (X) [دراسة تأثير زيادة عدد الحشرات على كمية الإنتاج لمحصول ما].

3- تغير ضغط دم الإنسان (Y) بتغير العمر (X) [دراسة تأثير العمر (X) على الضغط (Y)].

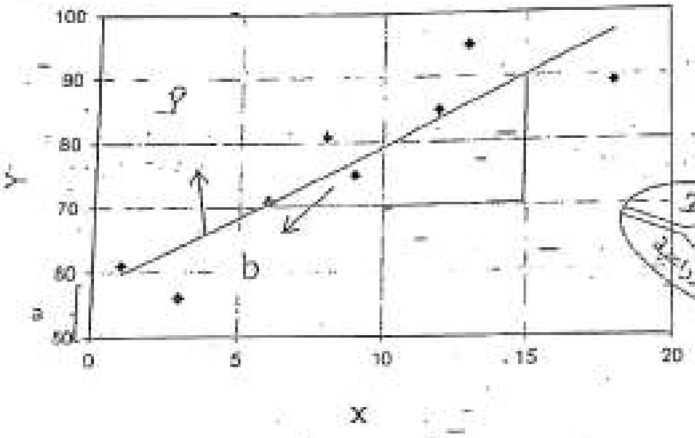
4- دراسة تأثير كمية السماد (X) على إنتاج الدونم (Y).

5- دراسة تأثير كمية البروتين (X) التي تتناولها الأبقار على الزيادة في وزنها (Y).

- **نموذج الانحدار الخطي البسيط Simple Linear Regression Model :**

إنّ دراسة العلاقة بين المتغيرين المستقل والتابع، تتم من خلال معادلة مستقيم والتي تدعى بمعادلة خط الانحدار، والتي تمثل العلاقة بين المتغيرين وتأخذ الشكل التالي: $\hat{Y}_i = \alpha + b\,x_i$ (٠)

- $\hat{Y}_i$: القيمة المقدّرة للمتغير التابع.
- x_i: المتغير المستقل.
- α: ثابت الانحدار، أو قيمة تقاطع خط الانحدار مع محور العينات.
- b: معامل الانحدار، أو قيمة ميل خط الانحدار.



شكل [illegible] يبين مستقيم خط الانحدار

- **معامل الانحدار The Regression Coefficient:**

هو مؤشر إحصائي يفسر مقدار التغير الحاصل في العامل التابع عندما يتغير العامل المستقل بمقدار وحدة واحدة.

يُحسب معامل الانحدار من العلاقة التالية:

$$b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

أو من العلاقة التالية:

$$b = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum {x_i}^2 - \frac{(\sum x_i)^2}{n}}$$

- $b > 0$: يكون تأثير x إيجابي على y. أكبر من صفر
- $b < 0$: يكون تأثير x سلبي على y. أصغر من صفر

- **ثابت الانحدار α:**

هو قيمة تقاطع خط الانحدار مع محور العينات، ويُحسب من العلاقة التالية:

$$\hat{y}_i = a + bx_i \qquad \alpha = \bar{y} - b\bar{x}$$

- وبالتعويض في العلاقة (•) نحصل على العلاقة التالية: $\hat{y}_i = \bar{y} + b(x_i - \bar{x})$ وتسمى هذه المعادلة بمعادلة التنبؤ التقديرية.

- **اختبار معنوية معامل الانحدار:**

النظرية الفرضية التي توضع تحت الاختبار هي (H0: معامل الانحدار لا يختلف عن الصفر).

$$H0: b = 0$$
$$H1: b \neq 0$$

وتحسب من خلال اختبار T بالعلاقة التالية:

$$T = \frac{b}{s\bar{b}}$$

- b: معامل الانحدار.
- sb:خطأ معادلة الانحدار أو الانحراف المعياري لمعامل الانحدار.
- $s\bar{b}$: الخطأ القياسي (المعياري) لمعامل الانحدار.

$$s\bar{b} = \frac{sb}{\sqrt{\sum(x_i - \bar{x})^2}}$$

$$sb = \sqrt{\frac{\sum\left(y_i - \hat{y}_i\right)^2}{n-2}}$$

- y_i: قيم العامل التابع (الأصلية) الحقيقية التي نتجت عن التجربة.
- $\hat{y}_i$: قيم العامل التابع التقديرية أو المتنبأ بها والتي نتجت من معادلة خط الانحدار.

- إذا كانت |T| المحسوبة > T الجدولية عند مستوى المعنوية ودرجات الحرية المحددة (n-2) ⟸ توجد فروق معنوية، وبالتالي معامل الانحدار يختلف عن الصفر ⟸ ترفض فرضية العدم وتقبل الفرضية البديلة، أي أنّ الانحدار معنوي، والمتغير المستقل له تأثير معنوي على الانحدار، ويمكن تمثيل العلاقة بين المتغير المستقل والتابع بخط مستقيم.
- إذا كانت |T| المحسوبة < T الجدولية عند مستوى المعنوية ودرجات الحرية المحددة (n-2) ⟸ لا توجد فروق معنوية، وبالتالي معامل الانحدار لا يختلف عن الصفر ⟸ تقبل فرضية العدم وترفض الفرضية البديلة، أي أن الانحدار غير معنوي، والمتغير المستقل ليس له تأثير معنوي

على الانحدار، والعلاقة بين المتغير المستقل والتابع ليست خطية، ويجب التفتيش عن شكل آخر للمعادلة يمثل البيانات بشكل فعلي.

مثال 1 :

في تجربة لدراسة العلاقة بين عمر النبات بالأسبوع وطوله بـ سم تم الحصول على النتائج التالية:

عمر النبات بالأسبوع x	1	2	3	4	5	6	7
طول النبات بـ سم y	5	13	16	23	33	38	40

المطلوب:

1- احسب معامل الارتباط بين عمر النبات وطوله وفسره.

2- اختبر معنوية معامل الارتباط وفسره.

3- احسب قيمة معامل التحديد وفسره.

4- أوجد معادلة خط الانحدار لطول النبات مع عمر النبات.

5- تنبأ بطول النبات عندما يكون عمر النبات (10 اسابيع).

6- مثل العلاقة بيانياً بخط مستقيم إذا كان ممكناً.

الحل:

x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
1	5	-3	-19	9	361	57
2	13	-2	-11	4	121	22
3	16	-1	-8	1	64	8
4	23	0	-1	0	1	0
5	33	1	9	1	81	9
6	38	2	14	4	196	28
7	40	3	16	9	256	48
$\sum x_i = 28$	$\sum y_i = 168$	0	0	28	1080	172

$$\bar{x} = \frac{\sum x_i}{n} = \frac{28}{7} = 4\ week$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{168}{7} = 24\ cm$$

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \rightarrow r = \frac{172}{\sqrt{28\times1080}} = 0.98$$

1- الارتباط طردي (موجب) وقوي جداً (تغير العمر يؤدي وبشكل قوي إلى تغير الطول).

2- اختبار معنوية معامل الارتباط:

- **طريقة استخراج r الجدولية:**

- يتم إنشاء خط شاقولي من مستوى المعنوية المطلوب.
- يتم إنشاء خط أفقي من درجات الحرية (n-2).
- يتقاطع الخطان عند رقم يمثل قيمة r الجدولية.
- يلاحظ أن قيمة r الجدولية عند مستوى 0.05 ودرجة حرية(5=2-7) = 0.754.
- يلاحظ أن قيمة r الجدولية عند مستوى 0.01 ودرجة حرية(5=2-7) = 0.874.

- يُلاحظ أن r المحسوبة > r الجدولية عند مستوى المعنوية 5% و 1% ودرجات الحرية5 ⟸ توجد علاقة معنوية جداً بين الصفتين، أي الارتباط بين الصفتين يختلف معنوياً عن الصفر، وأن هذا الاختلاف ليس راجع للصدفة.

قيم معامل الارتباط عند مستويات معنوية مختلفة

α / n	.1	.05	.02	.01	.001	α / n	.1	.05	.02	.01	.001
1	.98769	.99692	.999507	.999877	.9999988	16	.4000	.4683	.5428	.5897	.7084
2	.90000	.95000	.98000	.990000	.99900	17	.3887	.4555	.5285	.5751	.6932
3	.8054	.8783	.93433	.95873	.99116	18	.3783	.4438	.5155	.5614	.6787
4	.7293	.8114	.8822	.91720	.97406	19	.3687	.4329	.5034	.5487	.6652
5	.6694	.7545	.8329	.8745	.95074	20	.3598	.4227	.4921	.5368	.6524
6	.6215	.7067	.7887	.8343	.92493	25	.3233	.3809	.4451	.4869	.5974
7	.5822	.6664	.7498	.7977	.8982	30	.2960	.3494	.4093	.4487	.5541
8	.5494	.6319	.7155	.7646	.8721	35	.2746	.3246	.3810	.4182	.5189
9	.5214	.6021	.6851	.7348	.8471	40	.2573	.3044	.3578	.3932	.4896
10	.4973	.5760	.6581	.7079	.8233	45	.2428	.2875	.3384	.3721	.4648
11	.4762	.5529	.6339	.6835	.8010	50	.2306	.2732	.3218	.3541	.4433
12	.4575	.5324	.6120	.6614	.7800	60	.2108	.2500	.2948	.3248	.4078
13	.4409	.5139	.5923	.6411	.7603	70	.1954	.2319	.2737	.3017	.3799
14	.4259	.4973	.5742	.6226	.7420	80	.1829	.2172	.2565	.2830	.3568
15	.4124	.4821	.5577	.6055	.7246	90	.1726	.2050	.2422	.2673	.3375
						100	.1638	.1946	.2301	.2540	.3211

$$R_{square} = r^2 = (0.98)^2 = 0.96$$

3- 96% من الزيادة في طول النبات ناجم عن زيادة عمر النبات بالأسبوع و 4% من الزيادة في الطول يعود إلى عوامل أخرى غير متحكم بها(96% من التباين في الطول يُفسر من خلال العمر).

4- معادلة خط الانحدار من الشكل: $\hat{Y}_i = a + b\,x_i$ نحسب b و a

$$b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{172}{28} = 6.14$$

$$\alpha = \bar{y} - b\bar{x} = 24 - 6.14 \times 4 = -0.56$$

والمعادلة هي من الشكل : $\boxed{\hat{y}_i = -0.56 + 6.14\,x_i}$ إلا إنه يجب التأكد من معنوية معامل الانحدار حتى نستطيع أن نقول أن خط الانحدار مستقيم من الدرجة الأولى ويمكن رسمه.

يمكن كتابة المعادلة السابقة بالشكل التالي: $\hat{y}_i = \bar{y} + b(x_i - \bar{x})$ وبتعويض القيم المحسوبة يمكن كتابتها بالشكل التالي:

$$\hat{y}_i = 24 + 6.14(x_i - 4)$$

$$x = 1 \rightarrow \hat{y}_1 = 24 + 6.14(1 - 4) = 5.58\ cm$$
$$x = 2 \rightarrow \hat{y}_2 = 24 + 6.14(2 - 4) = 11.72\ cm$$

⋮

$$x = 7 \rightarrow \hat{y}_7 = 24 + 6.14(7 - 4) = 42.42\ cm$$

x_i	y_i	$\hat{y}_i$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$
1	5	5.58	-0.58	0.336
2	13	11.72	1.28	1.638
3	16	17.86	-1.86	3.460
4	23	24	-1	1
5	33	30.14	2.86	8.18
6	38	36.28	1.72	2.958
7	40	42.42	-2.42	5.856
$\sum x_i = 28$	$\sum y_i = 168$		0	23.60

- يُحسب الانحراف المعياري.
- يُحسب الخطأ المعياري.
- تُحسب قيمة T.

$$sb = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n - 2}} = \sqrt{\frac{23.60}{5}} = 2.17\ cm$$

$$s\bar{b} = \frac{sb}{\sqrt{\sum(x_i - \bar{x})^2}} = \frac{2.17}{\sqrt{28}} = 0.41\ cm$$

$$T = \frac{b}{s\bar{b}} = \frac{6.14}{0.41} = 14.97$$

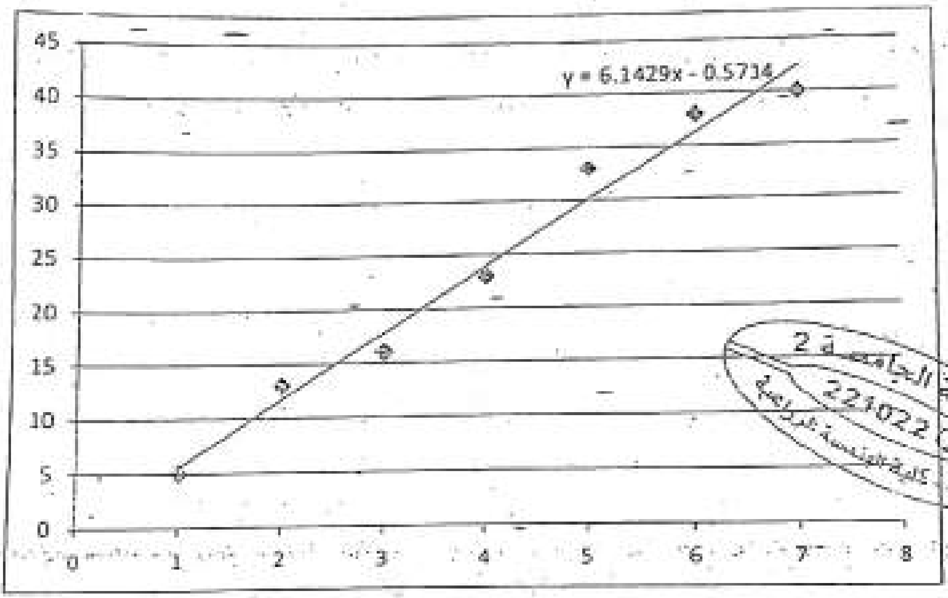
T: الجدولية عند مستوى معنوية 0.05 و درجة حرية 5 هي 2.571

- يلاحظ أن |T| المحسوبة > T الجدولية عند مستوى المعنوية ودرجة الحرية المحددة (n-2) ⟵ توجد فروق معنوية، وبالتالي معامل الانحدار معنوي ويختلف عن الصفر ⟵ تُرفض فرضية العدم وتُقبل الفرضية البديلة، أي أنَّ الانحدار معنوي، والمتغير المستقل له تأثير معنوي على الانحدار، وخط الانحدار مستقيم يمكن رسمه من الشكل:

$$\boxed{\hat{Y}_i = -0.56 + 6.14\, x_i}$$

- تفسر المعادلة بأنه لو زاد عمر النبات أسبوع واحد سوف يزداد طوله 6.14 سم.

5- طول النبات عندما يكون عمره 10 اسابيع: $\hat{Y} = -0.56 + 6.14 \times 10 = 60.84$



شكل (2) معادلة مستقيم الانحدار

مثال 2 :

البيانات التالية تمثل كمية الإنتاج لمحصول الشعير بـ كغ، والتي لوحظ أنها تزداد بزيادة كمية السماد المضافة إلى وحدة المساحة بـ كغ:

كمية السماد X	23	26	28	33	35	38	41
كمية الإنتاج Y	34	37	42	47	50	52	53

المطلوب:

- حساب قيمة معامل الانحدار.
- اختبر معنوية معامل الانحدار.
- ارسم الخط البياني إذا ثبتت معنويته .

<u>الجواب:</u> $b = 1.115, \alpha = 9.308$

المعدلة من الشكل: $\boxed{\hat{Y}_i = 9.308 + 1.115 x_i}$

- معامل الانحدار معنوي T= 11.357 وخط الانحدار مستقيم يمكن رسمه.

مثال 3 :

هناك علاقة تحكم بين كمية الإنتاج وطول السنبلة لمحصول القمح، ولبيان تلك العلاقة أخذت 10 مشاهدات وتم قياس طول السنبلة بـ سم وكمية الإنتاج بـ كغ وكانت النتائج كالتالي:

طول السنبلة x سم	5	7	8	9	10	11	12	15	16	17
كمية الإنتاج y كغ	4	6	7	7	8	10	11	12	12	13

المطلوب:

- هل يوجد تأثير معنوي لطول السنبلة على كمية المحصول.
- هل يمكن تمثيل العلاقة بيانياً بخط مستقيم.

<u>الجواب:</u> $b = 0.736, \alpha = 0.903$

- المعادلة من الشكل: $\boxed{\hat{Y}_i = 0.903 + 0.736 x_i}$
- معامل الانحدار معنوي T=12.536 وبالتالي يوجد انحدار معنوي لطول السنبلة على المحصول.
- خط الانحدار مستقيم يمكن رسمه.
- كل زيادة في طول السنبلة 1 سم سوف تزداد الغلة 0.736 كغ

الفصل الخامس

إعداد: د. قصي الهر

اختبار (ت)

T-Test of significance

يستخدم اختبار (t) للتعرف على مدى الاعتماد على نتائج التجارب التي تجرى في الميادين المختلفة سواء كانت مخبرية او حقلية ، ويرجع الفضل في اكتشاف هذا الاختبار الى العالم فيشر (Fisher) حيث اضاف الى الدراسات السابقة دراسة لتوزيع العينات الصغيرة حول متوسطها الحسابي بوحدات من الخطأ القياسي ووضع بهذا الشأن جداول (*) خاصة للتوزيع مبنية على اساس درجات الاحتمال وبنفس الوقت على درجات الحرية . ولقد سميت هذه الجداول باسم جداول (t) ويوجد بهذا الجدول قيم لـ (t) محسوبة للاحتمالات من ١٠ر. الى ٦ر. ولاى عدد من درجات الحرية .

ولاجراء الاختبار تحسب قيمة (ت) من البيانات المعطاة او المجمعة من التجربة كما يلي :

١- اذا كان متوسط المجتمع معروفا (u) يطبق القانون الاتي :

$$(ت) = \frac{\text{الفرق بين متوسط العينة والمجتمع}}{\text{الخطأ القياسي}}$$

أى بالرموز :

$$t = \frac{\bar{X} - u}{\sqrt{\frac{S}{n}}}$$

(*) [illegible] جداول [illegible] بالملحق في آخر هذا الكتاب

والنتيجة تسمى (t) المحسوبة التي تقارن بقيمة (t) الجدولية وعلــى المستوى للمعنوية المطلوب وفي العادة ما يكون هذا المستوى اما ١٪ او ٥٪ وتقرأ درجات الاحتمال من السطر المقابل لعدد درجات الحرية التي حسب منها الخطأ القياسي ، ويلاحظ من هذه المعادلة انه كلما قل الفرق بين متوسط العينة الحسابي ومتوسط المجتمع كلما صغرت قيمة (t) المحسوبة ونفس النتيجة فيما لو زادت قيمة الخطأ القياســي .

٢- اذا كان الاختبار خاص بمقارنة عينتين او معاملتين تصبح المعادلة كما يلي :

$$(ت) = \frac{\text{الفرق بين متوسطي العينتين}}{\text{الخطأ القياسي للفرق}}$$

أي بالرمــوز

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{d}}}$$

وبالطبع فان حساب الخطأ القياسي للفرق يختلف عن حساب الخطأ القياســي لعينة واحدة والذي سوف نوضحه فيما بعد .

الطريقــة العامة لاجراء اختبار (t) :

عند ما يراد تطبيق اخبار المعنوية (t) يؤخذ بعين الاعتبار الخطــوات التاليــة :

١- وضع النظرية الفرضية الاحصائية للتجربة او ما يطلق عليه "فرض العدم" علــى اساس انه لا يوجد فرق حقيقي بين متوسط التجربة ومتوسط المجتمع أو بيــن متوسطي العينتين او المعاملتين ، هذا وبنفس الوقت توضع النظرية البديلــة الا وهي انه هناك فروق . ويمكن التعبير عن هذه الخطوة بالرموز وفق مايلــي :

نظرية فرض العدم (مجتمع ،عينة)

$$u = \bar{X}$$

او

$$u - \bar{X} = 0$$

النظرية البديلة $\bar{X} \neq u$ (≠ لا يساوى)

نظرية فرض العدام $\bar{X}_1 = \bar{X}_2$

(عينتيــــن) $\bar{X}_1 - \bar{X}_2 = 0$

النظرية البديلة $\bar{X}_1 \neq \bar{X}_2$ (≠ لا يساوى)

٢- تقدر الاحصاءات اللازمة وهي المتوسط الحسابي للعينة والانحراف القياسي والخطــأ القياســي .

٣- تحسب قيمة (t) في ضوء المعادلات السابقة .

٤- تستخرج قيمة (t) من الجدول امام درجة الحرية التي حسب منها الخطأ القياسي عند درجتي احتمال ١٪ أو ٥٪ « الجدول موجود في آخر الفصل في الصفحة ٦٧،

٥- تقارن قيمة (t) المحسوبة من البيانات بقيمتها المستخرجة من الجــدول فان كانت :

آ- (t) المحسوبة اكبر من (t) الجدولية عند مستوى المعنوية ٥٪ اعتبر الفرق بين المتوسطين معنويا (Signiciant) أى حقيقــي وليس راجعا للصدفة ، وبذلك ترفض النظرية الفرضية .

ب- (t) المحسوبة اكبر من (t) الجدولية عند مستوى المعنوية ١٪ في هذه الحالة يعتبر أيضا الفرق بين المتوسطين معنوى جــــدا (Highly significant) وبذلك ترفض النظرية الفرضية .

ج- (t) المحسوبة اقل من (t) الجدولية معنى ذلك ان الفــرق بين المتوسطين غير معنوي (Insignificant) أى غير حقيقــي وغير مؤكد وراجع للصدفة ، وفي هذه الحالة تقبل النظرية الفرضية هذا ويستخدم اختبار (t) في عدد من المجالات يمكن حصرها فيما يلي :

١- لتقدير حدود متوسط المجتمع للعينات المدروسة .

٢- لمقارنة أزواج الافراد في التجارب المختلفة ذات البيانات غير المستقلة .

٣- لمقارنة المجموعات في التجارب المختلفة ذات البيانات المستقلة .

وسنبين فيما يلي كيفية التحليل الاحصائي في كل من المجالات السابقة وامكانيات الاستفادة منها في الوصول الى النتائج المستهدفة .

أولا : تقدير حدود متوسط المجتمع :

Estimate of population mean limits

لما كان متوسط المجتمع في اغلب الدراسات التي يجريها الباحثون غائبا وغير معروف نظرا لان العينات تكون هي اساس الدراسة لذا فان المتوسط الذى يحصل عليه من دراسة العينة لايمكن التعرف من خلاله على مساواته أو قربه او بعده عن متوسط المجتمع الا اذا تكررت التجربة من خلال أخذ عدد كبير من العينات، وهذا الامر يكون متعذرا جدا في عملية التطبيق . وللتغلب على هذه المشكلة وبهدف التعرف على متوسط المجتمع الذى اخذت منه العينة المدروسة يمكن الاستفادة من اختبار (t) في تقدير حدود متوسط المجتمع وهو ما يعبر عنه بالمدى الذى ينحصر فيه متوسط المجتمع او مجال الثقة ، ومجال الثقة هذا يتكون من حدين احدهما أدنى (L_1) والثاني أعلى (L_2) وطريقة الحصول على مجال أو حدود الثقة تعتمد أساسا على معادلة وجدول (t) كما يلي :

$$\pm t = \frac{\bar{x} - u}{s \bar{x}}$$

وبضرب الطرفين بالوسطين :

$$\bar{x} - u = \mp t \cdot s \bar{x}$$

وبنقل الحدود

$$u = \bar{x} \pm t \cdot s \bar{x}$$

ولدى استخراج قيمة (t) من الجدول عند درجة احتمال ٥ ٪ لعدد درجات

الحرية الذى حسب على اساسها الخطأ القياسي ومن ثم تعويضها في المعادلـــة فانه يمكن تحديد حدود الثقة (L_1 و L_2) لمتوسط المجتمع كما يلي :

الحد الأدنى

Lower limite $L_1 = \bar{X} - t05 \cdot S\bar{X}$

الحد الأعلى

upper limite $L_2 = \bar{X} + t05 \cdot S\bar{X}$

وعلى هذا الاساس فانه يمكن القول بان متوسط المجتمع يحتمل ان يقع بين القيمتيـــن (L_1 و L_2) في ٩٥ حالة من مائة ونكون على صواب ٩٥ ٪ من الحـــالات ، وقد نكون على خطأ في ٥ ٪ من الحالات ولكي نكون واضحين اكثر فاننا يمكن ان نقول بان مدى الثقة الذى نحصل عليه في هذه الحالة هو ٩٥ ٪ . هذا ويمكن ان نرفع حدود الثقة الى ٩٩ ٪ وذلك بالتعويض من قيمة (t) في المعادلة بقيمة (t) الجدولية عند درجة الاحتمال ١ ٪ وتحديد الحدود (L_1 , L_2) .

مثـال :

اذا كانت لديك الاحصاءات التالية لاحدى التجارب الحقلية الخاصة بمحصـول الذرة التي تهدف للوصول الى كمية المحصول في ستة من القطع المتساوية المزروعـة بمحصول الذرة والاحصاءات هي على التوالي :

متوسط محصول القطع = ١٥ كيلو غرام ($\bar{X}$)

الانحراف القياســي = ٢٫٢٧ كيلو غرام (S)

الخطأ القياســـي = ٠٫٩٢٤٨ كيلو غرام ($S\bar{X}$)

عدد القطـــــع = ٦ (N)

المطلوب تحديد حدود الثقة للمجتمع المأخوذة منه العينة السابقة وذلك علـى مستوى معنوية ٠٫٠٥ ٪

طريقة الحســاب :

طالما ان التجربة مختارة لعينة مؤلفة من ستة قطع من الحقل مزروعة بمحصـول الذرة ، فهذا يعني ان درجات الحرية تكون (٥) لذلك فاننا ننظر الى جـدول

(t) مقابل درجات الحرية رقم (٥) فنجد ان العدد على مستوى ٥ ٪ الموجود بالجدول هو (٢ر٥٧١) بعد ذلك نطبق القانون :

الحد الأعلى = $L_1 = \bar{X} + t05 \, . S \bar{X}$

بالتعويض ١٥ + (٢ر٥٧١ × ٠ر٩٤٨) = ١٧ر٤ كيلو غرام

الحد الأدنى = ١٥ - (٢ر٥٧١ × ٠ر٩٤٨) = ١٢ر٦ كيلوغرام

اذن يمكن القول بان متوسط المجتمع لقطع من الأرض التي توجد في حقل الذرة وتحت نفس الظروف سوف يقع بين ١٢ر٦ ، ١٧ر٤ كيلو غرام وهذا الاستنتاج يكون صحيحا في ٩٥ ٪ ، وطبعا يمكن التطبيق لاستخراج الثقة لدرجات احتمال ٩٩ ٪ اذ تعوض بالرقم المقابل لدرجات حرية (٥) ولدرجة احتمال ١ ٪ وفق نفس الاسلوب .

المقارنة تتم بين ازواج الافراد الذين يتشابهون بصفاتهم جميعا

ثانيا : اختبار (ت) ومقارنة ازواج الافراد Inaividual comparison

تعتبر طريقة مقارنة الأفراد التي توضع في ازواج في التجارب من ابسط واحسن الطرق الاحصائية في مجال تصميم التجارب . ولأجل توضيح أهمية هذا النوع من التحليل او التصميم نذكر المثال التالي :

في احدى تجارب التسميد التي تهدف الى الوصول أيهما أحسن وأفضل لمحصول القطن نترات النشادر أو نترات الجير . ولنفترض اننا زرعنا في ستة محطات منتشرة في انحاء الجمهورية في كل محطة قطعتين متساويتين ، احدهما سمدت بنترات النشادر والاخرى بنترات الجير ، فاذا قارنا الست قطع المزروعة بمحصول القطن والتي سمدة بنترات النشادر لمعرفة متوسط المحصول فيهما ، ومتوسط الست قطع الأخرى المزروعة والمسمدة بنترات الجير فان ذلك يعني ان المقارنة ستكون ليست فقط لتأثير السماد وانما يدخل فيها ضمنيا الفروق الناتجة لاختلاف التربة واختلاف الجو بين المحطات المختلفة بدلا أن يكون هذا الاختلاف ناتج عن

تأثير السماد فقط ، ولهذا يستعاض عن ذلك بمقارنة أزواج القطع المزروعة في نفس محطة التجارب والتي لها نفس نوع التربة ونفس الظروف الجوية ، وبذلك تكون واقعين اكثر ونزيد من دقة التجربة .

كما يمكن استخدام التصميم السابق أيضا في تجارب تغذية الحيوانات فاذا اردنا معرفة تأثير عليقتين مختلفتين على نسبة الدهن في الابقار او معدل النمو في التسمين للحيوان وجربنا لذلك العليقتين على مجموعتين متساويتين من الابقار فان النتائج التي نتوصل اليها لا تكون راجعة فقط الى تأثير تلك العلائق وانما قد ترجع الى الفروق التي تحدث للاختلافات التي توجد في تلك المجموعتين من الحيوانات سواء ما كان منها متعلقة بالتركيب الوراثي او العمر او تأثير البيئة ... الخ . ولهذا يفضل أن تستخدم مجموعة واحدة وتغذيتها على احدى العلائق ونأخذ النتائج . ثم بعدها نغذيها على العليقة الأخرى ونأخذ النتائج وبذلك نتغلب على الصعوبات الوارد ذكرها سابقا . هذا وقد يستخدم هذا التصميم في تجارب العقاقير الطبية وذلك باعطائها لنفس الشخص ومن ثم تقدير اثر ذلك قبل وبعد اخذ الدواء ... الخ .

وفي جميع مثل هذه التجارب او الحالات المدروسة تكون فيها النظرية الفرضية على الشكل الذى يقول بان الازواج التي يجرى عليها التجربة عبارة عن افراد من مجتمع طبيعي وضعت عشوائيا في ازواج . وعلى ذلك ينتظر ان تكون الفروق بين الازواج موزعة توزيعا طبيعيا حول متوسطها الحسابي الذى يساوى الصفر . وبهدف التعرف على طريقة التحليل الخاص بمقارنة أزواج الافراد واستخدام اختبار (ت) نفرض المثال التالي :

في إحدى تجارب زراعة الارز أجريت تجربة بهدف التعرف على نتائج المحصول فيما إذا زرع الأرز بطريقة الشتل ومن ثم بطريقة البذار وذلك لسبعة أصناف اختير لكل صنف مساحة متساوية في كل طريقة . وكانت النتائج حسب الجدول رقم (٦) التالي :

طريقة الزراعة / الصنف	أ	ب	جـ	د	هـ	و	ل
شتل	٢٥ر٣٠	٢١ر٧٠	٢٣ر٠٣	٢٠ر٢٢	٢٣ر٢٣	٢١ر٣٠	٢٣ر٣٢
بذار	٢١ر٣٠	٢٠ر٤٠	٢٠ر٤٣	٢٠ر٦٢	٢٠ر٣٣	١٨ر٣٠	٢٠ر٧٢

الجدول رقم (٦)

وبعد الحصول على نتائج هذه التجربة ، يكون الهدف الرئيسي هو معرفة افضل طريقة للزراعة بالنسبة لمحصول الارز وفقا للنتائج السابقة ، وهل طريقة الشتل احسن ام طريقة البذار . او بصورة اوضح نريد ان نعرف هل الفروق بين الطريقتين فروق مؤكدة احصائيا ام فروق ناتجة عن الصدفة ، علما بأن النظرية الفرضية (نظرية فرض العدم) في هذه الحالة تقول بأنه لا يوجد فرق بين الطريقتين وكلاهما يتبع مجتمع واحد وللوصول الى القرار النهائي لتلك الاستفسارات نقوم بتحليل هذه التجربة احصائيا مستخدمين اختبار (ت) لمقارنة ازواج الافراد . لذا ترتب النتائج في جدول رقم (٧) كما يلي علما بأننا سنرمز لطريقة الشتل بالرمز (X_1) ولطريقة البذار بالرقم (X_2) وللفرق بينهما بالرمز (D) وللمتوسط الحسابى للفرق بالرمز ($\bar{d}$) بعد وضع البيانات وفقا للجدول السابق مع اجراء عمليات الحساب للمؤشرات الاحصائية التي ستستخدم فى تقدير اختبار (ت) وهى :

١ـ حساب الفرق (D) بين محصول الشتل ومحصول البذار بالنسبة لكل صنف أى بين ($D = (X_1 - X_2$) ثم نقدر مجموع الفرق ($\sum D$) والذى يساوى :

$$\sum D = \sum_{i=1}^{n} (X_1 - X_2)$$

٢ـ حساب متوسط الفرق ($\bar{d}$) وذلك بقسمة مجموع الفروق ($\sum D$) على عدد الأصناف (n) .

$$\bar{d} = \frac{\sum D}{n} = ٢ر٣$$

الصنف	المحصول الشتل X_1	البذار X_2	الفرق للطريقتين $D = X_1 - X_2$	الانحراف عن متوسط الفرق $d = D - \bar{d}$	مربع الانحراف d^2	مربع الفرق $D^2 = (X_1 - X_2)^2$
١	٢٥ر٣٠	٢١ر٣٠	٤	١ر٧+	٢ر٨٩	١٦
٢	٢١ر٧٠	٢٠ر٤٠	١ر٣	١ر٠-	١ر٠٠	١ر٦٩
٣	٢٣ر٠٣	٢٠ر٣٣	٢ر٧	٠ر٤+	٠ر١٦	٧ر٢٩
٤	٢٠ر٢٢	٢٠ر٦٢	٠ر٤٠٠-	٢ر٧-	٧ر٢٩	٠ر١٦
٥	٢٣ر٢٣	٢٠ر٣٣	٢ر٩	٠ر٦+	٠ر٣٦	٨ر٤١
٦	٢١ر٣٠	١٨ر٣٠	٣	٠ر٧+	٠ر٤٩	٩ر٠٠
٧	٢٣ر٣٣	٢٠ر٧٢	٢ر٦	٠ر٣+	٠ر٠٩	٦ر٧٦
المجموع	$\sum X_1 =$ 158.1	$\sum X_2 =$ 142	$\sum D =$ 16.1	$\sum \bar{d} = 0$	$\sum d^2$ 12.28	$\sum D^2$ 49.31
المتوسط	$\bar{X}_1 =$ 22.5857	$\bar{X}_2 =$ 20.2867	$\bar{d} =$ 2.3			

الجدول رقم (٧)

ويمكن الحصول على متوسط الفرق ايضا وذلك بطرح متوسط طريقة البذار من متوسط طريقة الشتل :

$$\bar{d} = \bar{x}_2 - \bar{x}_1 = \text{٢ر٣}$$

ويلاحظ اننا رمزنا لمتوسط الفرق ($\bar{d}$) تميزاً له عن المتوسط الحسابي ($\bar{X}$) .

٣- حساب انحراف (d) كل فرق من الفروق (D) عن متوسط الفرق ($\bar{d}$) وذلك بالطرح :

$$d = D - \bar{d}$$

ويلاحظ ان مجموع الانحرافات عن متوسط الفرق d يساوى صفر حيث تمثل

هذه احد الصفات الاساسية للمتوسط الحسابي .

٤- حساب مربع الانحرافات عن متوسط الفرق ويتم ذلك عن طريق تربيع الانحرافات عن متوسط الفرق وبذلك نحصل على d^2 وبعد ان يتم جمعها ولكل القيم نحصل على مجموع مربعات الانحرافات عن متوسط الفرق $\sum d^2$ وهو يساوي في مثالنا (١٢ر٢٨) أي :

$$\sum d^2 = \sum (D - \bar{d})^2 = \text{١٢ر٢٨}$$

٥- حساب الانحراف القياسي (S) وهو يساوي الجذر التربيعي لحاصل قسمة مجموع مربعات الانحرافات على عدد درجات الحرية أي على عدد الفروق ناقصا واحد كما في المعادلة التالية :

$$S = \sqrt{\frac{\sum (D-\bar{d})^2}{n-1}} = \sqrt{\frac{\sum d^2}{n-1}}$$

$$S = \sqrt{\frac{\text{١٢ر٢٨}}{\text{٦}}} = \text{١ر٤٣}$$

٦- يقدر الخطأ القياسي للفرق ويرمز له بالرمز ($S\bar{d}$) وهو يساوي

$$S\bar{d} = \sqrt{\frac{S^2}{n}} = \frac{S}{\sqrt{n}}$$

$$= \frac{\text{١ر٤٣}}{\sqrt{\text{٧}}} = \text{٠ر٥٤١}$$

٧- بحسب قيمة (ت) وهي وفقا للمعادلة تساوى :

$$\mp t = \frac{\bar{X}_1 - \bar{X}_2}{S\bar{d}} = \frac{\bar{d}}{S\bar{d}}$$

أى ان مقدار (ت) المحسوبة بعد التطبيق بالارقام هى :

$$t = \frac{٢٠ر٢٨ - ٢٢ر٥٨٧}{٠ر٥٤١}$$

$$\frac{٢ر٣}{٠ر٥٤١} = ٤ر٢٥$$

٨- يقارن قيمة (ت) المحسوبة بقيمتها الجدولية من أجل فحص النظرية الفرضية والتى تقول بأنه لا يوجد فرق بين طريقتى الزراعة ولهذا نستخرج قيمة (ت) الجدولية ولعدد درجات الحرية التى استخرج منها الانحراف القياسي وهى فى التجربة تساوى (٧-١) أى ستة وبالنظر الى جدول (ت) عند درجات الحرية (٦) نرى أن درجات الاحتمال هى :

$$t_{0.5} = ٢ر٤٤٧$$

$$t_{0.1} = ٣ر٧٠٧$$

وهذا يعنى أنه يوجد فرق حقيقي ومؤكد إحصائيا لطريقتى الزراعة الشتل والبذار وهذا الفرق معنوى جدا وترفض النظرية الفرضية والتى تقول بانه لا يوجد فروق بين طريقتى الزراعة حيث أن طريقة الزراعة شتلا قد زاد عن طريقة الزراعة بذارا بالنسبة لمحصول الأرز وهذه الزيادة معنوية احصائيا ومؤكدة فى ٩٩ ٪ من الحالات .

٩- يمكن الوصول الى حساب قيمة الانحراف القياسي ومن ثم الخطأ القياسي بطريقة أخرى خلاف المتبعة وتدعى بطريقة تربيع القيم او الطريقة المباشرة . ويترك شرحها للدروس العملية منعا للخلط والغموض .

١٠- اذا اردنا الحصول على معامل الاختلاف بين طريقتى الزراعة شتلا وبذارا فاننا نطبق القانون لمعامل الاختلاف وهو :

$$C.V. = \frac{S}{\bar{X}G} . 100$$

حيث :

C.V = معامل الاختلاف Coeficent of varation

S = الانحراف القياسي

$\bar{X}$ G = المتوسط العام

بالتطبيق معامل الاختلاف يساوى :

$$\frac{١٫٤٣}{٢١٫٤٣٥٧} \times ١٠٠ = ٦٫٦ \ \%$$

ثالثا : اختبار (ت) لمقارنة المجموعات Group comparison

فى الحالات التى لا يستطيع الباحث وضع الأفراد فى ازواج لمقارنتها ببعضها البعض نظرا لعدم وجود ارتباط بينهما او استقلال البيانات ، يلجأ الباحث الى تقسيم الافراد للتجربة الى مجموعتين وبشكل عشوائي يتم اختيار احد المجموعات ومعاملتها بمعاملة ما ومن ثم يتم مقارنتها بالمجموعة الاخرى ، وهنا تتم المقارنة بين الفرق بين متوسطى المجموعتين ولا تكون المقارنة بين ازواج الافراد .

هذا والنظرية الفرضية الاحصائية فى مثل هذا التصميم والتي تتخذ كأساس فى هذا النوع من التجارب تقول ان الفرق بين متوسطى العينتين يرجع الى الاخطاء العشوائية ، ولذلك يستعمل اختبار (ت) لمعرفة هل المتوسطين مأخوذين من مجتمعين مختلفين أم من مجتمع واحد وفى اطار ذلك يمكن ان تكون المجموعتين متساويتين في عدد الافراد او تكون غير متساوية في عدد الافراد وبالطبع يختلف التحليل فى كل حالة وفقاً لما يلى :

آ- الحالة التي يكون فيها عدد الأفراد متساوى في المجموعتين فان اختبار (ت) يجرى وفقا للمثال التوضيحي التالي :

فى تجربة لمعرفة اثر المبيدات الحشرية على انتاج محصول القطن ، زرعت عشرون قطعة بمحصول القطن من صنف معين ثم اختير منها عشر قطع تركت للطبيعة

دون استخدام المبيد ، اما العشر قطع الاخرى فقد رشت بالمبيد ، وعند إنتاج المحصول جمع القطن من عدد معين من الخطوط في كل قطعة بالنسبة للمحصول المعامل ولنفس العدد من الخطوط للمحصول غير المعامل وكانت كمية القطن بالكيلو غرام مبينة بالجدول رقم (٨) حسبما يلي :

طريقة الانحراف

طريقة تربيع القيم

رقم القطعة	غير المعامل				المعامل			
	محصول	الانحراف	مربع الانحراف	مربع القيم	محصول القطعة	الانحراف	مربع الانحراف	مربع القيم
	X_1	$x_1=X_1-\bar{X}_1$	x_1^2	X_1^2	X_2	$x_2=X_2-\bar{X}_2$	x_2^2	X_2^2
١	٣ر١	١-	١	٩٦١	٦ر١	٠ر١-	٠ر٠١	٣٧ر٢١
٢	٣ر٨	٠ر٣-	٠ر٠٩	١٤ر٤٤	٥ر٦	٠ر٦-	٠ر٣٦	٣١ر٣٦
٣	٤ر٣	٠ر٢+	٠ر٠٤	١٨ر٤٩	٦ر٤	٠ر٢+	٠ر٠٤	٤٠ر٩٦
٤	٣ر٩	٠ر٢-	٠ر٠٤	١٥ر٢١	٥ر٨	٠ر٤-	٠ر١٦	٣٣ر٦٤
٥	٤ر٠	٠ر١-	٠ر٠١	١٦	٦	٠ر٢-	٠ر٠٤	٣٦
٦	٤ر٠	٠ر١-	٠ر٠١	١٦	٥ر٧	٠ر٥-	٠ر٢٥	٣٢ر٤٩
٧	٤ر٥	٠ر٤+	٠ر١٦	٢٠ر٢٥	٥ر٩	٠ر٣-	٠ر٠٩	٣٤ر٨١
٨	٤ر٢	٠ر١+	٠ر٠١	١٧ر٦٤	٧ر٥	١ر٣+	١ر٦٩	٥٦ر٢٥
٩	٤ر٣	٠ر٢٠	٠ر٠٤	١٨ر٤٩	٧	٠ر٨+	٠ر٦٤	٤٩
١٠	٤ر٩	٠ر٨+	٠ر٦٤	٢٤ر٠١	٦	٠ر٢-	٠ر٠٤	٣٦
	41	٥	2,04	170.14	62	0	3.22	387.72
المجموع	$\sum X_1=\sum X_1=$ ٤١	$\sum x_1=$	$\sum x_1^2=$	$\sum X_1^2=$	$\sum X_2=$ ٦٢	$\sum x_2=$	$\sum x_2^2=$	$\sum X_2^2=$
	٤ر١	صفر	٢ر٠٤	١٧٠ر١٤	$\bar{X}_2=$ ٦ر٢	صفر	٣ر٣٢	٣٨٧ر٧٢

الجدول رقم (٨)

وفي هذا المثال نتبع الخطوات التالية في التحليل بهدف الوصول الى الهدف المطلوب :

١- نقدر المتوسط الحسابي للمحصول المعامل وغير المعامل وفقا لما يلي :

$$\bar{X}_1 = \frac{\sum X_1}{n_1} = \frac{\text{٤١}}{\text{١٠}} = \text{٤ر١}$$

$$\bar{X}_2 = \frac{\sum X_2}{n_2} = \frac{\text{٦٢}}{\text{١٠}} = \text{٦ر٢}$$

٢- نقدر مجموع مربعات الانحرافات عن المتوسط الحسابي لكل من المحصـــــول للمعامل وغير المعامل وفقا لما يلي :

أ- مجموع المربعات للانحرافات لغير المعامل =

$$SS_1 = \sum (X_1 - \bar{X}_1)^2 = \text{٢ر٠٤}$$

أو بالطريقة المباشرة (تربيع القيم)

$$SS_1 = \sum X_1^2 - \frac{(\sum X)^2}{n} = \text{١٧٠ر١٤} - \frac{\text{(٤١)}^2}{\text{١٠}}$$

$$= \text{٢ر٠٤}$$

ب- مجموع المربعات للانحرافات للمعامل =

$$SS_2 = \sum (X_2 - \bar{X}_2)^2 = \text{٣ر٣٢}$$

أو بالطريقة المباشرة (تربيع القيم)

$$SS_2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n} =$$

$$\text{٣٨٧ر٧٢} - \frac{\text{(٦٢)}^2}{\text{١٠}} = \text{٣ر٣٢}$$

٣- نقدر المجموع المشترك لمربعات الانحرافات للمعاملتين :

$$SS_p = SS_1 + SS_2$$ المجموع المشترك لمربعات الانحرافات

$$= \text{٢ر٠٤} + \text{٣ر٣٢}$$

$$= \text{٥ر٣٦}$$

٤- تحسب متوسط المجموع المشترك لمربعات الانحرافات او ما يطلق عليه بالتباين المشترك وذلك بقسمة المجموع المشترك لمربعات الانحرافات للمعاملتين مقسوما على عدد درجات الحرية للمعاملتين وهى تساوى: =

$$(n_1 + n_2 - 2)$$

وبما ان

$$n_1 = n_2$$

اذن تصبح

$$(2n - 2)$$

أو

$$2 (n-1)$$

ssp و ١

اذن متوسط المجموع المشترك (التباين المشترك) يساوى

$$S^2_p = \frac{\sum x_1^2 + \sum x_2^2}{n_1 + n_2 - 2} = \frac{\text{٥ر٣٦}}{\text{٢-١٠+١٠}} = \text{٠ر٢٩٧٨}$$

٥- يقدر الخطأ القياسي للفرق بين المتوسطين وهو يساوى :

$$S_{\bar{d}} = \sqrt{\frac{S^2_p}{n_1} + \frac{S^2_p}{n_2}}$$

وحيث ان التباين (S^2_p ، S^2_p) متساوى (التباين المشترك) اذن يمكن وضع المعادلة بالشكل التالى علما بأن عدد الافراد متساوى بالمعاملتين

$$S_{\bar{d}} = \sqrt{\frac{2\ S_p^2}{n}}$$ = الخطأ القياسي للفرق

وبعد التطبيق تكون قيمة الخطأ القياسي للفرق هى :

$$\text{٠ر٢٤٤} = \sqrt{\frac{\text{٠ر٢٩٧٨×٢}}{\text{١٠}}} = \sqrt{\frac{\text{٠ر٢٩٧٨}}{\text{١٠}} + \frac{\text{٠ر٢٩٧٨}}{\text{١٠}}} =$$

٦- نحسب قيمة (ت) وفقا للمعادلة التالية :

$$\mp t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{d}}}$$

$$\frac{٤ر١ - ٦ر٢}{٠ر٢٤٤} = - ٨ر٦١$$

٧- باستخراج قيمة (ت) من الجدول عند درجات الحرية التي استخرج منها الخطأ القياسي وهي (٢٠-٢) = ١٨ نجد ان الارقام الواردة في الجدول هي :

t_{05} = ٢ر١٠١

t_{01} = ٢ر٨٧٨

أي ان (ت) المحسوبة اكبر من (ت) الجدولية لكلا المستويين وهذا يعني أن النظرية الفرضية مرفوضة أي ان المعاملتين ليسا من مجتمع واحد لا نه توجد فروق بين المعاملتين مؤكدة احصائيا وهي تشكل فروق حقيقية وليست راجعة الى الصدفة وهذه الفروق هي فروق معنوية جداً اى بمعنى آخر فان الرش بالمبيدات قد أفاد في زيادة محصول القطن بصفة معنوية جدا .

٨- يمكن تقدير معامل الاختلاف في هذه التجربة بتطبيق القانون التالي :

$$C . V \% = \frac{S}{\bar{X} G} . 100$$

حيث ان :

C.V. = معامل الاختلاف

S = الانحراف القياسي او جذر التباين المشترك (S_p^2)

$\bar{X}$ G = المتوسط العام

اذن معامل الاختلاف = $\frac{\sqrt{٠ر٢٩٧٨}}{٥ر١٥} \times ١٠٠$

= ١٠ر٦ ٪

اختبار (ت) لمقارنة مجموعتين مختلفتين في عدد افراد كل منهما :

في كثير من التجارب قد يكون عدد الافراد في المجموعتين قيد المقارنة غير

متساويين ويرجع السبب فى ذلك إما لتعذر ايجاد عدد متساوى من الافراد لمقارنتهما معا او لاننا قد نبدأ التجربة بعدد متساوى للمجموعتين المراد مقارنتهما ببعض. ولكن لاسباب خارجية قد ينقص افراد احد المجموعتين عن الاخرى او العكس ومثال ذلك موت بعض حيوانات التجارب الامر الذى يتطلب من الباحث استبعاد بعض الافراد. وعلى هذا الاساس تصبح عدد الافراد فى المجموعتين غير متساوى وفى هذه الحالة لا تختلف الاسس الاحصائية التي بنيت عليها طريقة اختبار (ت) فى تحليل النتائج المتحصل عليها في المجموعتين التي اعدت للمقارنة وانما تعدل طريقة التحليل وطريقة الحساب تعديلا طفيفا فقط بحيث يؤخذ بعين الاعتبار ان عدد افراد المجموعة الاولى لايساوى عدد افراد المجموعة الثانية وتطبيق ذلك اثناء حساب الوسط الحسابى او الانحراف القياسي او الخطأ القياسي للفرق .. الخ .

مثال :

فى إحدى تجارب وزارة الزراعة للتعرف على بعض الصفات المدروسة بين نوعين من الدجاج البلدى والدجاج اللجهورن أخذت وزن البيضة من كلا النوعين كأحد الصفات المدروسة وفيما يلي بعض الاوزان بالغرام لعدد من البيض للنوع الاول جدول رقم (٩) والثانى جدول رقم (١٠) والمطلوب هل يوجد بين النوعين من البيض فروق حقيقية ؟ .

الجدول رقم (٩) (لجهورن)

n_1	الوزن (غ) X_1	الانحراف عن المتوسط $x_1 = X_1 - \bar{X}_1$	مربع الانحراف $(X_1 - \bar{X}_1)^2$	مربع الوزن X_1^2
١	٥٦	٢	٤	٣١٣٦
٢	٥٢	٢-	٤	٢٧٠٤
٣	٥٠	٤-	١٦	٢٥٠٠
٤	٤٧	٧-	٤٩	٢٢٠٩
٥	٤٦	٨-	٦٤	٢١١٦
٦	٥٧	٣	٩	٣٢٤٩

تابع جدول رقم (٩)

٧	٥٢	٢	٤	٢٧٠٤
٨	٥٨	٤	١٦	٣٣٦٤
٩	٦٥	١١	١٢١	٤٢٢٥
١٠	٥٦	٢	٤	٣١٣٦
١١	٥٦	٢	٤	٣١٣٦
١٢	٥٣	١-	١	٢٨٠٩
١٣	٥٦	٢	٤	٣١٣٦
١٤	٥٢	٢-	٤	٢٧٠٤
المجموع	∴ $\sum X$ =٧٥٦		٣٠٤	٤١١٢٨

$\bar{X}_1$ = ٥٤

الجدول رقم (١٠) (البلـــــد ى)

n_2	الوزن (غ) X_2	الانحراف من المتوسط $X_2 - \bar{X}_2$	مربع الانحراف $(X_2-\bar{X}_2)^2$	مربع الوزن X_2^2
١	٢٤	٤-	١٦	١١٥٦
٢	٣٦	٢-	٤	١٢٩٦
٣	٣٦	٢-	٤	١٢٩٦
٤	٣٥	٣-	٩	١٢٢٥
٥	٣٣	٥-	٢٥	١٠٨٩
٦	٤٢	٤	١٦	١٧٦٤
٧	٤١	٣	٩	١٦٨١
٨	٣٩	١	١	١٥٢١
٩	٤٤	٦	٣٦	١٩٣٦
١٠	٤٢	٤	١٦	١٧٦٤
١١	٣٣	٥-	٢٥	١٠٨٩
١٢	٣٩	١	١	١٥٢١

تابع جدول رقم (١٠)

١٢	٣٧	-١	١	١٣٦٩
١٤	٣٦	-٢	٤	١٧٩٦
١٥	٤٣	٥	٢٥	١٨٤٩
	$\sum X_2=$			$X_2^2=$
المجموع	٥٧٠	.	١٩٢	٢١٨٥٢
	$\bar{X}_2$ ٣٨			

طريقة الحساب :

١- يحسب المتوسط الحسابى لوزن البيضة لكل من المجموعتين وذلك بجمع الاوزان لكل صنف وتقسيمه على عدد البيض .

المتوسط الحسابى لوزن بيضة اللجهورن $\bar{X}_1 = \frac{\sum X_1}{n_1} = \frac{٧٥٦}{١٤} = ٥٤$

المتوسط الحسابى لوزن بيضة البلدى $\bar{X}_2 = \frac{\sum X_2}{n_2} = \frac{٥٧٠}{١٥} = ٣٨$

٢- تقدير مجموع مربعات الانحرافات لكلا النوعين من البيض نبدأها باللجهورن :

$$SS_1 = \sum (X_1-\bar{X}_1)^2 = ٣٠٤$$

أو

$$= \sum X_1^2 - \frac{(\sum X_1)^2}{n_1}$$

$$= ٤١١٢٨ - \frac{(٧٥٦)^2}{} = ٣٠٤$$

مجموع مربعات الانحرافات للبيض البلدى :

$$SS_2 = \sum (X_2-\bar{X}_2)^2 = ١٩٢$$

أو

$$= \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}$$

بالتعويض $٢١٨٥٢ - \frac{(٥٧٠)^2}{١٠} = ١٩٢$

٣- نقدر المجموع المشترك لمربعات الانحرافات في كلا النوعين (المجموعتين) .

$$SS_p = SS_1 + SS_2$$

$$٤٩٦ = ٣٠٤ + ١٩٢$$

٤- نقدر متوسط المجموع المشترك لمربعات الانحرافات او التباين المشترك (S^2_p) الذى يساوى .

$$S_p^2 = \frac{SS_1 + SS_2}{n_1 + n_2 - 2}$$

وبعد التعويض $\frac{١٩٢ + ٣٠٤}{١٤+١٥-٢} = ١٨ر٣٧$

٥- نقدر الخطأ القياسي للفرق وهو في هذه الحالة :

$$S_{\bar{d}} = \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

وبعد التعويض $\sqrt{\frac{١٨ر٣٧}{١٥} + \frac{١٨ر٣٧}{١٤}} = ١ر٥٩٣$

٦- تحسب قيمة (ت) وفق القانون التالي :

$$\mp t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{d}}}$$

$$\frac{٣٨ - ٥٤}{١ر٥٩٣} = ١٠ر٠٠٤$$

٧- بمقارنة قيمة (ت) المحسوبة مع قيمتها الجدولية على مستوى ٥ ٪ ، ١ ٪.ولما كانت درجات الحرية التي استخرج منها الخطأ القياسي وهى (١٤+١٥-٢ = ٢٧ درجة حرية وبالرجوع الى الجدول نرى ان القيم الجدولية هـي

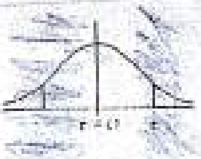
٢ر٠٥٢ ، ٢ر٧٧١ على التوالي .

٨- وحيث ان قيمة (ت) المحسوبة اكبر من قيمتها الجدولية فانه يمكن القبول بانه يوجد فرق معنوى جدا بين وزن بيضة كل من اللجهورن والبلدى أو يمكن القول بان مجموعتى البيض ليستا من مجتمع واحد وبذلك نرفض النظرية الفرضية الاحصائية التى تقول بأنه لا يوجد فرق بين النوعين من البيض وعلى هذا الاساس فان بيض النوع الاول اللجهورن يتفوق على نوع البيض الثاني البلدى بدرجة موءكدة احصائيا التي تنتج عن فروق حقيقية وليست ناتجة تلك الفروق عن الصدفة .

جدول (III) توزيع t

Student's t-Distribution

A denotes the sum of the two tail areas for the values of t given below. ν denotes the number of degrees of freedom (df).



ν or df	A = 0.1	A = 0.05	A = 0.02	A = 0.01	A = 0.001
1	6.314	12.706	31.821	63.657	636.619
2	2.920	4.303	6.965	9.925	31.598
3	2.353	3.182	4.541	5.841	12.941
4	2.132	2.776	3.747	4.604	8.610
5	2.015	2.571	3.365	4.032	6.859
6	1.943	2.447	3.143	3.707	5.959
7	1.895	2.365	2.998	3.499	5.405
8	1.860	2.306	2.896	3.355	5.041
9	1.833	2.262	2.821	3.250	4.781
10	1.812	2.228	2.764	3.169	4.587
11	1.796	2.201	2.718	3.106	4.437
12	1.782	2.179	2.681	3.055	4.318
13	1.771	2.160	2.650	3.012	4.221
14	1.761	2.145	2.624	2.977	4.140
15	1.753	2.131	2.602	2.947	4.073
16	1.746	2.120	2.583	2.921	4.015
17	1.740	2.110	2.567	2.898	3.965
18	1.734	2.101	2.552	2.878	3.922
19	1.729	2.093	2.539	2.861	3.883
20	1.725	2.086	2.528	2.845	3.850
21	1.721	2.080	2.518	2.831	3.819
22	1.717	2.074	2.508	2.819	3.792
23	1.714	2.069	2.500	2.807	3.767
24	1.711	2.064	2.492	2.797	3.745
25	1.708	2.060	2.485	2.787	3.725
26	1.706	2.056	2.479	2.779	3.707
27	1.703	2.052	2.473	2.771	3.690
28	1.701	2.048	2.467	2.763	3.674
29	1.699	2.045	2.462	2.756	3.659
30	1.697	2.042	2.457	2.750	3.646
40	1.684	2.021	2.423	2.704	3.551
60	1.671	2.000	2.390	2.660	3.460
120	1.658	1.980	2.358	2.617	3.373
∞	1.645	1.960	2.326	2.576	3.291

Table II is abridged from Table III of Fisher and Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, published by Oliver and Boyd Limited, Edinburgh, by permission of the authors and publishers.

الفصل السادس

إعداد : د. قصي الجمر

- اختبار كا٢ (مربع كـــاى) -

(x^2) Chi squar

تقسم البيانات التي نحصل عليها من الابحاث والتجارب الى قسمين رئيسيين وذلك من وجهات النظر الاحصائية ويطلق على القسم الأول بالقياسات ومثال ذلك الوزن والطول وكمية المحصول وغير ذلك من الصفات الكميـــة .

أما النوع الآخر من البيانات فيطلق عليه التعداد الذى يصف عدد القياسات أو التكرارات أو عدد النباتات او عدد الحشرات . . . الخ وتختلف طرق التحليل الاحصائي للبيانات الكمية (جميع الاختبارات التي مرت معنا) عن التحليل الاحصائي للبيانات العددية ، وسنتناول في هذا الجزء من المقرر اجراء اختبار المعنوية للبيانات العددية .

يعتبر مربع كاى أو (x^2) لحسن المطابقة او ما يطلق عليه احيانا إختبار التطابق النسبي من أهم الطرق الرئيسية التي تستخدم في تحليل الصفات العددية ، ويرجع اكتشاف هذا الاختبار الى العالم " كارل بيرسون " ويقوم اساس هذا الاختبار على مقارنة مجموعة من النتائج المشاهدة او المتحصل عليها من تجربة حقيقية بمجموعة أخرى فرضية وضعت على اساس النظرية الفرضية التي يراد إختبارها والتحقق منها .

والنظرية الفرضية التي توضع تحت الاختبار هي أنه لا يوجد فروق بين القيم المشاهدة او الفعلية التي نتجت من التجربة وبين القيم النظرية أو المتوقعة واختبار كاى مربع يوضح مدى انطباق القيم أو التكرارات المشاهدة على تلك النظريـــة .

• وقيمة مربع كاى عبارة عن مجموع مربعات الفروق بين القيم المشاهدة او الفعلية وبين القيم النظرية او المتوقعة مقسوما على القيم النظرية او المتوقعة وذلك كما في المعادلة التالية

$$x^2 = \sum_{1}^{k} \left[\frac{(O_i - E)^2}{E} \right]$$

حيث أن :

x^2 = قيمة مربع كاى

k = عدد الفئات او المجموعات لاعداد الافراد أو المشاهدات

O = القيم الفعلية او المشاهدة Observed

E = القيم المتوقعة او النظرية Expected

علما بأن هذا المؤشر الاحصائي (كا^{٢}) يتبع في سلوكه توزيع كا^{٢} حيث يوجد جدول خاص * بهذا التوزيع ولعدد درجات حرية ($k-1$) حيث تقارن قيمة كا^{٢} الجدولية بقيمتها المحسوبة والتي على اساس تلك المقارنة نرفض أو نقبل بالنظرية الفرضية ولكي يكون هذا المؤشر قريبا بقدر الامكان من التوزيع النظرى الموجود بالجدول فانه يجب أن تكون حجم العينة كبيرا بحيث لا يقل العدد الكلي للتكرارات عن ٥٠ وان تكون تكرارات الفئة الواحدة على الاقل خمسة أو عشرة ويكون ذلك أحسن وتعطي النتائج بشكل أفضل .

مثال لاستخدام مربع كاى لدرجة حرية واحدة :

تفيد التعليمات الموجودة في نشرة بشأن تأثير مبيد حشرى على قتل نوع معين من الحشرات بأن مفعوله يؤثر بصورة قاتلة على ٨٠ ٪ من عدد الحشرات الحية والاختبار ما اذا كان هذا المبيد ينطبق على الظروف العملية ، اجريت تجربة حقيقية رش فيها المبيد على نباتات تحتوى على عدد معروف من الحشرات (٤٠٠)

* الجدول موجود في آخر الفصل في الصفحة ٧٣ ،

حشرة ثم أحصي عدد الحشرات الحية فوجد عددها (١٠٠) بينما الميتة كانت (٣٠٠) والمطلوب معرفة صحة إنطباق التعليمات على الواقع العملي :

الحـــــل :

ان هذه البيانات تعتبر بيانات عددية وللتوصل الى المطلوب يجرى اختبار كا٢ ولذا ترتب البيانات في جدول تفريغ رقم (١٧) من اجل الحصول على المؤشرات التي تلزمنا لاستخراج قيمة كا٢ وفقاً لما يلـــي :

جـــــدول (١٧)

المشاهدة	تكرار مشاهد O	تكرار نظرى E	$O - E$	$\frac{(O-E)^2}{E}$
حشرات حية	١٠٠	٨٠	٢٠	$\frac{(٢٠)^٢}{٨٠} = ٥$
حشرات ميتة	٣٠٠	٣٢٠	-٢٠	$\frac{(-٢٠)^٢}{٣٢٠} = ١ر٢٥$
المجمـــوع	٤٠٠	٤٠٠	صفـر	٦ر٢٥

٤٠٠ × ٠ر٢ = ٨٠

٤٠٠ × ٠ر٨ = ٣٢٠

بتطبيق المعادلة لاستخراج قيمة كا٢

$$X^2 = \sum_{1}^{k} \frac{(O-E)^2}{E}$$

$$\frac{(١٠٠-٨٠)^٢}{٨٠} + \frac{(٣٠٠-٣٢٠)^٢}{٣٢٠} = ٦ر٢٥$$

في هذه الحالة فان (k) تساوى لفئتين أو مجموعتين المجموعة الاولى تمثل الحشرات الحية والمجموعة الثانية تمثل الحشرات الميتة وعلى ذلك فـــان درجات الحرية تصبح (٢-١) = ١ أى درجة حرية واحدة وبالنظر الى جدول توزيـــع

(كا٢ وعند درجة حرية واحد نجد أن :

$$\overset{2}{X}o1 = ٦٫٦٤$$

$$\overset{2}{X}o5 = ٣٫٨٤$$

أى أن كا٢ المحسوبة اكبر من كا٢ الجدولية عند مستوى معنوية ٥ ٪ وعلى هذا الاساس فاننا نرفض النظرية الفرضية أى أن الحالة العملية للمبيد لا تتفق أولا تتطابق مع القانون النظرى له الذى يقول بموت ٨٠ ٪ من الحشرات بينما وجد فرق حقيقي ومؤكد احصائيا مما يدل على رفض النظرية الفرضية ، علماً بأنه لو كانت قيمة كا٢ المحسوبة أقل من قيمتها الجدولية دل ذلك على أنه لا توجد فروق حقيقية او معنوية وتقبل النظرية الفرضية .

مثال لاستخدام مربع كاى لاكثر من درجة حرية واحدة :

لا يختلف التحليل عما سبق الا بدرجات الحرية حيث كان المثال السابق يتعرض في التحليل الى مجموعتين من المكرارات لذا كانت درجات الحرية هي واحد واذا كان التحليل لأكثر من مجموعتين من المكررات فان درجات الحرية تكون وفقا لعددها منقوصا منها واحد كما هو الحال في المثال التالي :

وجد باحث أن احدى الصفات الوراثية يتحكم بها عاملين وراثين سائدان هما A , B وان النسل في الجيل الثاني يتوزع بالنسبة التالية ٩ : ٣ : ٣ : ١ قام باحث آخر بفحص ٥٥٦ حالة من حالات الجيل الثاني ووجد ان النسل يتوزع فيها وفقا لاربعة مجموعات هي على التوالي ٣١٥ ، ١٠١ ، ١٠٨ ، ٣٢ ، والمطلوب إختبار مدى تمشي التوزيع الشاهد مع التوزيع النظرى السابق .

الحل : للوصول الى المطلوب توضع البيانات والنتائج في جدول رقم (١٨)

جدول رقم (١٨)

المشاهدة	تكرار مشاهد O	تكرار نظرى E	الفرق O - E	$\frac{(O-E)^2}{E}$	طريقة حساب E
AB	٣١٥	٣١٢ر٧٥	٢ر٢٥	٠ر٠١٦	$556 \times \frac{9}{16}$
Ab	١٠١	١٠٤ر٢٥	- ٣ر٢٥	٠ر٠١	$556 \times \frac{3}{16}$
aB	١٠٨	١٠٤ر٢٥	٣ر٧٥	٠ر١٣٥	$556 \times \frac{3}{16}$
ab	٣٢	٣٤ر٧٥	- ٢ر٧٥	٠ر٢١٨	$556 \times \frac{1}{16}$
مجموع	٥٥٦	٥٥٦	صفر	٠ر٤٧٠	

اذا :

$$\text{٠ر٤٧} = \frac{(\text{٢ر٢٥})^{٢}}{\text{٣١٢ر٧٥}} + \frac{(\text{٣ر٢٥٠})^{٢}}{\text{١٠٤ر٢٥}} + \frac{(\text{٣ر٧٥})^{٢}}{\text{١٠٤ر٢٥}} + \frac{(\text{٢ر٧٥} -)^{٢}}{\text{٣٤ر٧٥}}$$

وبمقارنة كا^{٢} المحسوبة بـ كا^{٢} الجدولية وعلى درجات حرية (٤ - ١) = ٣

نجد أن $X^{2}01 = \text{١١ر٣٤}$

$X^{2}05 = \text{٧ر٨٢}$

وهي اكبر من القيمة المحسوبة لذا فاننا نقبل النظرية الفرضية نظرا لعدم وجود فروق معنوية مؤكدة احصائيا وهذا يعني ان الصفة الوراثية تتوزع في الجيل الثاني في اطار التوزيع النظرى الخاص وهي النسبة ٩ : ٣ : ٣ : ١ .

تابع جدول (IV)

TABLE IV. Chi-Square Distribution
(continued)

ν or df	A = 0.30	A = 0.20	A = 0.10	A = 0.05	A = 0.02	A = 0.01	A = 0.001
1	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	4.88	5.99	7.78	9.49	11.67	13.28	18.46
5	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	7.23	8.56	10.64	12.59	15.03	16.81	22.46
7	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	9.52	11.03	13.36	15.51	18.17	20.09	26.12
9	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	11.78	13.44	15.99	18.31	21.16	23.21	29.59
11	12.90	14.63	17.28	19.68	22.62	24.72	31.26
12	14.01	15.81	18.55	21.03	24.05	26.22	32.91
13	15.12	16.98	19.81	22.36	25.47	27.69	34.53
14	16.22	18.15	21.06	23.68	26.87	29.14	36.12
15	17.32	19.31	22.31	25.00	28.26	30.58	37.70
16	18.42	20.46	23.54	26.30	29.63	32.00	39.25
17	19.51	21.62	24.77	27.59	31.00	33.41	40.79
18	20.60	22.76	25.99	28.87	32.35	34.80	42.31
19	21.69	23.90	27.20	30.14	33.69	36.19	43.82
20	22.78	25.04	28.41	31.41	35.02	37.57	45.32
21	23.86	26.17	29.62	32.67	36.34	38.93	46.80
22	24.94	27.30	30.81	33.92	37.66	30.29	48.27
23	26.02	28.43	32.01	35.17	38.97	41.64	49.73
24	27.10	29.55	33.20	36.42	40.27	42.98	51.18
25	28.17	30.68	34.38	37.65	41.57	44.31	52.62
26	29.25	31.80	35.56	38.88	42.86	45.64	54.05
27	30.32	32.91	36.74	40.11	44.14	46.96	55.48
28	31.39	34.03	37.92	41.34	45.42	48.28	56.89
29	32.46	35.14	39.09	42.56	46.69	49.59	58.30
30	33.53	36.25	40.26	43.77	47.96	50.89	59.70

Table IV is abridged from Table IV of Fisher and Yates: *Statistical Tables for Biological, Agricultural, and Medical Research*, published by Oliver and Boyd Limited, Edinburgh, by permission of the authors and publishers.

الفصل السابع

تصميم التجارب Design of experiments:

مقدمة:

إن الوظائف الرئيسية للعلم science هو:-

١- الوصف Description

٢- التحليل والتفسير Analysis and interpretation

٣- التنبؤ Expectation or prediction

٤- التحكم Controlling:

ويستطيع العلم أن يؤدي هذه الوظائف عن طريق البحث العلمي

ويعرف البحث العلمي بأنه:" العملية التي يتم من خلالها الاستفسار عن طبيعة حدث معين وأسبابه والنتائج التي تترتب عليه."

وتعتبر التجربة Experiment من أهم وسائل البحث العلمي، فالتجربة هي أساس المعرفة، إذ إنها أداة الطريقة العلمية للوصول إلى معرفة حقيقة الأشياء التي نهتم بها في جميع أوجه النشاط الإنساني (أثر مواعيد الزراعة على إنتاجية محصول معين،تأثير البرامج التلفزيونية في رفع وعي المشاهدين، أثر عقار معين على التخفيف من الإصابة بمرض معين ، تأثير نسبة الذكاء على رفع مستوى التحصيل الدراسي للطلبة) ويتم الوصول إلى المعرفة عن طريق المشاهدة وجمع البيانات وتحليلها وتفسيرها، وهذا يؤدي إلى استخلاص أكبر قدر من المعلومات تفيدنا في معرفة حقيقة الظواهر و إمكانية التنبؤ بها و التحكم فيها.

ولقد ساهم علم الإحصاء و الإحصائيين في تقدم البحث العلمي ،عن طريق إيجاد العديد من التصميمات التجريبية،بالإضافة إلى الأساليب التحليلية الملائمة لها، ويعتبر العالم الانكليزي Fisher في مقدمة العلماء الذين ساهموا في تطوير تصميم التجارب،حتى أصبح يكون فرعا مستقلا و هاما من فروع الإحصاء،وقد وضعت أول التصميمات في محطة تجارب زراعية في انجلترا،وأصبحت تستخدم حاليا في معظم ميادين البحث العلمي كالطب والهندسة وغيرها.

وللتعريف بتصميم التجارب نستطيع أن نقول بأنه:

(فرع الإحصاء الذي يهتم بتطبيق الطريقة الإحصائية في التجربة العلمية،ويشتمل هذا الفرع على التعريف بالتصميمات المختلفة،وطريقة تنفيذها وتحليل بياناتها،للوصول إلى قرارات علمية بدرجة كافية من الدقة،وبأقل تكلفة ممكنة.)

وقبل الدخول في أساسيات تصميم التجارب،لابد لنا من التعرض لبعض المصطلحات والتي سيتردد ذكرها في كافة أجزاء هذا المقرر،وهذه المصطلحات هي:

التجربة:Experiment

تعتبر التجربة من أهم وسائل الطريقة العلمية للبحث،وتعرف بأنها:

"دراسة اختباريه جرى تصميمها بصورة مسبقة للحصول على حقائق جديدة،أو الإثبات أو نفي معلومات سابقة."

وتجرى التجارب الزراعية في الحقول و معامل الأبحاث والغرف الزجاجية،بهدف دراسة تأثير بعض المعاملات الزراعية على محصول معين،وذلك بهدف الوصول إلى توصيات تؤدي إلى زيادة الإنتاج الزراعي.

ولتوضيح مفهوم التجربة نورد المثال التالي:

أراد باحث مقارنة إنتاجية صنف معين من القمح في حالة التسميد و عدم التسميد (طن/هكتار)

وقد اختار قطعتين متشابهتين من الأرض،واستخدم السماد في إحداها،ولم يستخدمه في الأخرى،وتفترض أنه كرر هذه التجربة خمس مرات،وحصل على النتائج التالية:

إنتاجية القمح بدون سماد "طن/هكتار" Without fertilizer	إنتاجية القمح مع السماد "طن/هكتار" With fertilizer
2.5	3.3
2.3	3.7
2.0	3.5
2.2	[illegible]
[illegible]	2.8
[illegible]	متوسط الانتاجية مع السماد X1 = 3.4

تتلخص أهداف هذه التجربة في نقطتين أساسيتين،وهما:

1- هل هناك فرق في الإنتاجية بين وضع السماد من عدمه،أو بعبارة أخرى هل الفرق في المحصول هو فرق معنوي أم ناشئ بمحض الصدفة، وللإجابة على السؤال يقع اختيار الفرض الصفري أو فرض العدم (Null hypothesis) وهو $H_0: u_1 = u_2$ ، u_1 ، u_2 هما متوسطا إنتاجية القمح،ويقول هذا الفرض بأنهما متساويان،أي أنه ليس هناك فرقا معنويا في إنتاجية القمح في حالة التسميد و عدم التسميد،وهذا ضد الفرض البديل (Alternative hypothesis) وهو: $H_a: u_1 \neq u_2$ والذي يقر بأن هناك فرقا معنويا بين الحالتين.

2- تقدير قيمة الفرق:ويكون هذا التقدير مرتبطا بالسؤال الأول،أي إذا استنتجنا عن طريق الاختبار أن هناك فرقا معنويا،فيصبح من باب الاستدلال تقدير قيمة ذلك عن طريق معرفة حدود الثقة (confidence limits) لمتوسطات المجتمعات المدروسة.

وتعمل معظم التجارب لتحقيق الهدفين التاليين:

1- اختبار نظريات فرضية Testing hypothesis

2- تقدير فروق المعاملات Estimating treatment differences

والتجربة أيضا هي: مجموعة من الإجراءات تستخدم لأخذ عينات عشوائية من مجتمعات البحث، ففي مثالنا السابق هناك مجتمعان:

المجتمع الأول: ويتمثل في مجموعة القيم التي تدل على إنتاجية ذلك الصنف من القمح تحت ظروف عدم التسميد،ويتكون هذا المجتمع من قيم لا تحصى ولا تعد، ولذلك يسمى مجتمعا لا نهائيا، ولدراسة هذا المجتمع سحبت عينة مكونة من 5 مشاهدات وهي تلك التكرارات الخمسة.

المجتمع الثاني: إنتاجية القمح في حالة التسميد، وله نفس خصائص المجتمع الأول.

والتكرارات الخمسة لكلتا الحالتين ما هي إلا عينات أخذت من مجتمعي البحث، لكي تكون الوسيلة العملية لدراستها،ووضع استدلالات أو استنتاجات حولها، وهذا هو الأمر السائد في الإحصاء الاستدلالي (Inferential statistics) وهو سحب عينة للاستدلال بها عن مجتمع، فاللجوء إلى المعاينة العشوائية هدفه توفير المال والوقت والجهد في إجراء البحوث، بالإضافة إلى عدم إمكانية شمول البحث لجميع عناصر المجتمع الإحصائي المدروس في الكثير من الحالات، ومن هنا كانت أهمية اللجوء إلى سحب العينات العشوائية الممثلة للمجتمع على حقيقته، من أجل التعرف على خصائص العينة، ومن ثم تعميم نتائجها على المجتمع الإحصائي المدروس.

المعاملة Treatment

هي مجموعة من الظروف وضعت تحت سيطرة الباحث ليتمكن من قياس تأثيرها على صفة محددة لمواد التجربة (محصول معين ـ حيوان معين) فإذا أردنا معرفة أثر التسميد بسماد معين على محصول الذرة ، فإننا نزرع صنف معين مع تسميد جزء منه بالسماد ، وترك جزء أخر بدون تسميد ، مع توحيد كافة الظروف الأخرى في التجربة ،فالمثال يحتوي على معاملتين (التسميد وعدم التسميد)

مثال أخر: في إحدى التجارب الحقلية لاختبار أثر مواعيد الزراعة على إنتاج محصول القطن صممت تجربة بحيث اختير أربع مواعيد للزراعة (مبكر A ، متوسط B ، متأخر C ، متأخر جدا D)

وقد تم زراعة /16/ قطعة تجريبية ، بحيث أن كل /4/ قطع وزعت في إحدى المواعيد السابقة.

فكل موعد من المواعيد الأربعة يعتبر معاملة.

العامل: Factor

الاصطلاح عامل يتشابه في المعنى مع الاصطلاح معاملة ، ولو أن معناه أوسع ، فهو: نوع من المعاملة التي تحتوي على تقسيمات متعددة تسمى المستويات Levels

والمستويات نوعان ، هما :

١) مستويات وصفية أو نوعية ، فمثلا عامل التربة قد يتكون من ثلاثة مستويات: (رملية – سلتية – طينية)

٢) مستويات كمية ، فمثلا عامل الرطوبة قد يتكون من أربعة مستويات: 10% - 20% – 30% - 40%

وتعتبر التجربة الخاصة بالأمثلة الواردة في الفقرة الخاصة بالمعاملة هي تجربة بسيطة أو تجربة ذات عامل واحد (Simple or single Factor experimnt)

لأن جميع الظروف موحدة ما عدا التسميد أو موعد الزراعة فقط. وقد يرغب الباحث في دراسة تأثير التسميد على خمسة أصناف من الذرة ، فيلجأ إلى زراعة كل من الأصناف الخمسة ، مع تسميد جزء من المساحة المخصصة لكل صنف ، وترك الجزء الأخر بدون تسميد ، مع تثبيت كل الظروف والمعاملات الزراعية الأخرى ، وعلى ذلك فهذه التجربة تحتوي على عاملين : الأول هو التسميد وله مستويين والثاني هو الأصناف ويتكون من خمسة مستويات ، وتسمى هذه التجربة بالتجربة العاملية (Factorial experiment)

الوحدة التجريبية : Expermintal unit or plot

وهي أصغر وحدة أو قسم في التجربة ، ففي التجارب الحقلية فإن الوحدة التجريبية هي عبارة عن قطعة أرض عوملت كوحدة من حيث معاملات التجربة ، فإذا أردنا مقارنة خمسة أصناف من القطن ، وكل منها في خمس قطع من الأرض (خمس مكررات) ، فإننا نقسم الحقل إلى /25/ قطعة متساوية من الأرض ، بحيث يوزع كل من الأصناف في خمس قطع متساوية وموزعة عشوائيا ، وكل من هذه القطع يطلق عليها وحدة تجريبية.

المعاملة القياسية أو معاملة الشاهد: Control or check treatment

وهي إحدى معاملات التجربة التي لا يكون الباحث في حاجة إلى إثبات أهميتها ، وإنما يدخلها في تجاربه لمقارنة معاملات التجربة المختلفة بها. ووجود معاملة الشاهد ضروري وشرط أساسي لزيادة دقة التجربة، لأن اختبار دقة المعاملات وتفسير النتائج يكون على أساس المقارنة معها ، وتعامل معاملة الشاهد بنفس معاملات التجربة ، فيما عدا المعاملات تحت الاختبار.

فمثلا: لاختبار إنبات بذور القطن في ظروف ملحية مختلفة ، حضرت أربعة تراكيز من الأملاح (0.5 ، 1 ، 1,5 ، 2 ميليموز) وعوملت واحدة بالماء المقطر . فالماء المقطر هو معاملة الشاهد.

ومثلا: لدراسة إنتاجية أربعة أصناف مستوردة من الخيار ، بالمقارنة مع الصنف المحلي . فالصنف المحلي هو معاملة الشاهد.

الخطأ التجريبي: Experimental Error

هو التباين بين الوحدات التي طبقت عليها نفس المعاملة ،وبما إن الاختلاف هو من خصائص الظواهر الحيوية،فيتكون الخطأ التجريبي من مجموعة العوامل غير المتحكم فيها،والكامنة داخل المادة التجريبية ولهذا عندما نتحدث عن الاختلاف بين الوحدات التجريبية أو الخطأ التجريبي،فلا يعني ذلك أنه حصلت أخطاء في التجربة،وإنما ذلك نتيجة الاختلاف في المادة التجريبية،مثل الاختلاف في التربة ،أو في الحيوانات التي أجريت عليها التجربة،أو في طريقة الحصاد أو الزراعة ....الخ.

وتتلخص مصادر الخطأ التجريبي في نقطتين أساسيتين:

١- عدم تجانس الوحدات التجريبية(الاختلاف في خصوبة التربة، التركيب الوراثي للحيوانات أو النباتات،أو العوامل الجوية......الخ)

٢- طريقة تنفيذ التجربة(اختلاف الأشخاص الذين يقومون بإجراء إعمال التجربة فيما بينهم في قوة الإبصار أو أداء العمليات أو قوتهم الجسمانية)

والهدف من التصميمات المختلفة هو إن كل تصميم متقدم يتميز عن التصميم الذي سبقه بقدرته بشكل أكبر على تصغير قيمة الخطأ التجريبي،من خلال التحكم في الوحدات التجريبية وتأمين تجانسها ، وذلك لأنه كلما كان الخطأ التجريبي صغيرا كانت التجربة أدق.

أساسيات تصميم التجارب:

يشترط في التصميمات الحديثة إن تعطي تقديرا للخطأ التجريبي مع امكانية تقليله وأن يكون بالإمكان القيام بالاختبارات والتقديرات المطلوبة في البحث ، ووضعت أساسيات تصميم التجارب لتوفر تلك المطالب، وهي ثلاثة قواعد : التكرار- التعشية- التحكم في الوحدات التجريبية أو مبدأ وحدة الاختلاف.

١- مبدأ وحدة الاختلاف: **unity of variance principle** :

ويقصد بذلك المساواة التامة في معاملة كافة القطع أو الوحدات التجريبية، أي توحيد كافة شروط التجريب باستثناء العامل المدروس، بحيث يكون الاختلاف الوحيد بين المعاملات هو العامل المدروس فمثلا عند تصميم تجربة لدراسة تأثير خمسة تركيزات من مبيد حشري معين على محصول الذرة الصفراء فيجب إن تكون المعاملات الزراعية المختلفة بالنسبة للمحصول المعامل بهذه التركيزات واحدة من مواعيد وطرق زراعة، ري ،تسميد، عزيق......الخ، والاختلاف الوحيد بين هذه المعاملات هو تركيز المبيد الحشري فقط.

٢-التوزيع العشوائي أو التعشية: **randomization:**

ويقصد بذلك إن يتم توزيع المعاملات المطلوب دراستها على وحدات التجربة بشكل عشوائي، دون أي تحيز، إي دون إن تؤثر العوامل الذاتية أو اختلاف الزمان والمكان على عملية توزيع المعاملات، ويعتبر استعمال التعشية بمثابة عملية التأمين ضد التحيز بحيث تمنح كل وحدة تجريبية نفس الفرصة لاستلام أية معاملة في التجربة.

ويختلف التوزيع العشوائي للمعاملات أو الوحدات التجريبية حسب تصميم التجربة المستعمل ،ويجرى التوزيع باستخدام جداول الأرقام العشوائية أو القرعة وغيرها

٣-التكرار: **replication**

ويقصد بالتكرار معاملة أكثر من وحدة تجريبية واحدة على الأقل بنفس المعاملة، وللتكرار فوائد متعددة وهي:

أ-إيجاد إمكانية تقدير الخطأ التجريبي .

ب-تقليل الخطأ التجريبي .

ج-تقديم استنتاج أعم وأشمل عن التجربة حيث يمكن تكرار نفس التجربة في نفس العام، أو أعوام متتالية، أو في أماكن مختلفة

بعد استعراض المبادئ الأساسية في تصميم التجارب فلا بد لنا من القول بان الهدف الأساس من تصميم التجارب هو ترتيب الوحدات التجريبية في التجربة (grouping of plotes) إي توزيع المعاملات على الوحدات التجريبية بطريقة تضمن الحصول على اقل خطأ ممكن ،وفي نفس الوقت تقدير غير متحيز لأثر العوامل المراد دراستها على هذه الوحدات

فعندما تكون المادة التجريبية متجانسة (حقل متجانس- حيوانات متجانسة مثلا) ففي هذه الحالة توزع المعاملات عشوائيا على الوحدات التجريبية

ويكون تصميم التجربة بسيطا وهو التصميم العشوائي الكامل (CRD)

وقد يكون التصميم أكثر تعقيدا في حالة عدم التجانس بحيث يؤدي إلى تجميع المعاملات في قطاعات كاملة complete blocks

(تصميم القطاعات العشوائية الكاملة RCBD- المربع اللاتيني LSD) أوفي قطاعات ناقصة incomplete blocks

(تصميم القطع المنشقة SBD)

الخطوات المنطقية للبحث العلمي :

تتلخص هذه الخطوات بما يلي :

١- تحديد مشكلة البحث والأهداف والفرضيات المراد اختبارها والتقديرات الإحصائية المطلوبة، وتحديد التحليل الإحصائي المناسب لتحقيق تلك الأهداف وإعطاء إجابة وافية ومستفيضة عن كل الأسئلة المطروحة .

٢- اختيار المعاملات: أي اختيار المعاملات أو العوامل ومستوياتها التي سيقع بحثها في التجربة.

٣- اختيار الصفة أو الصفات المدروسة :تشتمل هذه الخطوة على تحديد الصفة أو الصفات المدروسة، أي القياسات المطلوب تسجيلها، والتي تعطي معلومات كافية حول مشكلة البحث، وغالبا ما تعرف الصفة المدروسة بالمتغير التابع أو الاستجابة ويرمز لها بالرمز y .

٤- تصميم التجربة: أي إيجاد التصميم المناسب للدراسة التي تحت البحث، بحيث تحدد طريقة التعشية الملائمة، وعدد التكرارات (حجم العينة) وكل ذلك في ضوء أهداف البحث، وكمية المواد التجريبية المتاحة والميزانية المخصصة لتلك التجربة، مع العلم إن أحسن التجارب هي أبسطها تصميما وأيسرها تكلفة ، وأسهلها تحليلا وتفسيرا.

٥- تنفيذ التجربة: وهنا يتطلب أن يشرف الباحث بنفسه على تنفيذ التجربة. ومن أكثر المشاكل شيوعا في هذه المرحلة :سوء استعمال التعشية – عدم الدقة في اخذ القياسات للصفات المدروسة – عدم توحيد طريقة التعامل مع الوحدات التجريبية في التجربة

٦- تحليل البيانات : mintab-sas-spss-genstat-mstatc . وهي مجموعة من البرامج الاحصائية.

٧- النتائج: وهي عصارة الجهد كله، فبعد التحليل تستخلص الاستنتاجات للأسئلة المطروحة والتي على ضوئها توضع التوصيات المناسبة، وهناك طرق علمية لعرض نتائج التجربة في تقارير واضحة، بحيث تمكن غير المتخصصين من فهم واستيعاب ما توصل إليه البحث من نتائج نذكر منها : تجنب استخدام الألفاظ الإحصائية الصعبة – تلخيص المتوسطات في جداول واضحة – توفير الرسوم البيانية كلما أمكن ذلك.

من كل ما تقدم من موضوعات في هذا الفصل ، فإننا نستخلص النقاط التالية:

١-الهدف الرئيسي لأي تصميم هو استخراج تأثير المعاملات، وفصلها عن أثر الاختلافات الموجودة أساسا بين الوحدات التجريبية.

٢- يتوقف اختيار تصميم دون أخر على المتطلبات التالية:

أ- أن يكون التصميم بسيطا وسهلا في التحليل

ب- اختيار التصميم الذي يعطي اكبر دقة ممكنة، بأقل تكاليف ممكنة.

ج- عدم اختلاف الوحدات التجريبية بطريقة منتظمة (التعشية)

د- أن يمكن هذا التصميم من تقدير الخطأ التجريبي وتقليله

ه- سهولة تحليل النتائج في حالة فقد محصول وحدة تجريبية أو أكثر

٣- هناك عدة طرق من شأنها إن تحسن من دقة التجربة، ونذكر منها الأتي :

أ- تكبير حجم التجربة بزيادة عدد التكرارات

ب- المزيد من العناية في اختيار المعاملات

ج- تحسين طريقة تنفيذ التجربة

د- تقسيم الوحدات التجريبية على قطاعات متجانسة

اختبار F وتحليل التباين F-Test and analysis of variance

إعداد أ.م. مصطفى العمر

إن اختبار t يستخدم لاختبار معنوية الفرق بين متوسطي عينتين A,B (عينات غير مستقلة في أزواج – عينات مستقلة لا توجد في أزواج)، ولكن في الواقع هناك تجارب كثيرة ومسائل إحصائية تتعلق باختبار متوسطات ثلاثة مجتمعات فأكثر، وفي هذه الحالة تصبح عملية التحليل وفق اختبار t أكثر تعقيدا.

فعلى سبيل المثال: إذا كان لدينا أربع أنواع من المبيدات، ونريد معرفة ما إذا كانت المبيدات الأربعة تعطي نفس النتائج أم أن هناك اختلافا معنويا بينها؟

والجواب على هذا السؤال يمكن تحقيقه بواسطة اختبار t ، ولكن بشرط مقارنة متوسطي مجتمعين لكل زوج من المجتمعات الأربعة، وهذا يعني القيام بستة اختبارات ل t ، فإذا كانت متوسطات المجتمعات الأربعة هي: u_4، u_3، u_2، u_1 ،

فإن الفرضيات المطلوب اختبارها هي:

$H_0:u_1=u_4$

$H_0:u_3=u_4$

$H_0:u_1=u_3$

$H_0:u_2=u_4$

$H_0:u_1=u_2$

$H_0:u_2=u_3$

والعيب الرئيسي لهذه الطريقة هو أنها ليست عملية، حيث يزداد عدد المقارنات بسرعة كلما ازداد عدد المجتمعات، ويمكن حساب عدد المقارنات المحتملة من خلال المعادلة التالية:

$$\text{Number of comparisons} = \frac{T(T-1)}{2}$$

T: عدد المجتمعات أو عدد المعاملات أو عدد المتوسطات.

فعلى سبيل المثال: إذا كان عدد المتوسطات 6 فهذا يعني أننا نحتاج للقيام باختبار t 15 مرة.

وقد تمكن العالم الإنكليزي Fisher أن يضع أسس توزيع يتمكن فيه الباحث من إجراء اختبار واحد لتحليل نتائج التجارب التي يزيد فيها عدد المعاملات أو المجموعات عن معاملتين، وأطلق عليه توزيع F تكريما له، وقد طوره العالم Snedecor من خلال حساب جداول خاصة لتوزيع F تتميز برقمين لدرجات الحرية، أحدهما أفقي للمتغير المستقل أو بين المعاملات والأخر عمودي للخطأ التجريبي.

وقد أدى اكتشاف هذه الطريقة إلى تقدم سريع في علم الإحصاء وتصميم التجارب، وأصبحت هذه الطريقة أداة فعالة في تحليل الأبحاث الحديثة.

وبالعودة إلى تجربة المبيدات الأربعة، فإنه يمكننا أن نختبر فرضية عدم واحدة بدلا من ست فرضيات، وهي : $H_0:U_1=U_2=U_3=U_4$

أي هل المتوسطات الأربعة متساوية أم لا، وعند رفض هذا الفرض، فإن الفرض البديل يكون على الشكل التالي:

$u_1 \neq u_2 \neq u_3 \neq u_4$ متوسط واحد ui على الأقل يختلف عن الباقي : Hə

والاختبار المناسب لفرض العدم هو اختبار F الذي نحصل عليه باستخدام طريقة تحليل التباين، ويعتبر تحليل التباين من أقوى الطرق الإحصائية المتوفرة في علم الإحصاء، وأكثرها شيوعا في البحوث التجريبية.

وتتلخص طريقة تحليل التباين في عملية حسابية تجرى على البيانات، وذلك بتجزئة مجموع مربعات الانحرافات الكلية للاستجابة $\Sigma(X_{ij} - \bar{X})^2$ إلى عدد من المجاميع المختلفة طبقا للمصادر المسببة للاختلاف، والتي تختلف حسب التصميم المستخدم في التجربة، ومن هنا فإن تحليل التباين يفي بالهدف الأول للتجارب العلمية وهو:
هل هناك فروق بين المتوسطات ؟ (Testing hypothesis)

ويعتبر التصميم العشوائي الكامل (The completely Randomized design CRD) :
أبسط صور التقسيم التي تستخدم في تصميم التجارب ، حيث تقسم فيه المصادر المسببة للاختلاف أو الانحرافات، أو المجموع الكلي لمربعات الانحرافات (SSO) إلى مصدرين رئيسيين هما:

١ - الاختلافات التي ترجع إلى مجموعات أو معاملات التجربة، ويطلق عليه الاختلافات بين المعاملات Between treatment ويرمز له بالرمز SST

٢ - الاختلافات التي توجد بين أفراد التجربة داخل المعاملات، ويطلق عليها الاختلافات داخل المعاملات Within treatment أو خطأ العينات Randon error أو الخطأ التجريبي Experiment error أو البواقي Remainder ويرمز له sse

ومصادر هذا الخطأ هي:

أ- الاختلافات الطبيعية بين الأفراد.

ب- العجز عن التحكم في كل الظروف المحيطة بالتجربة.

ج- العجز عن توحيد الطرق المستخدمة في التجربة.

ويتم حساب sso و sst و sse على الشكل التالي:

١) في البداية نقوم بحساب معامل التصحيح Correction Factor CF على الشكل التالي:

$$CF = \frac{X..^2}{tr} \quad \text{أو} \quad CF = \frac{G^2}{N}$$

G أو ..X: المجموع الكلي للقيم الناتجة بالتجربة

t: عدد المعاملات

r: عدد الأفراد داخل كل معاملة، أي تكرار المعاملة.

$\left.\begin{matrix} t \\ r \end{matrix}\right\} N$

٢) ثم نقوم بحساب sso على الشكل التالي:

أ- الطريقة المباشرة: $SSO = \sum_{i=1}^{t}\sum_{j=1}^{r}(Xij - \bar{X}..)^2$

$\bar{X}..$: المتوسط العام للتجربة أي المجموع الكلي لقيم التجربة على العدد الكلي لقيم التجربة.

Xij : قيمة وحدة تجريبية في معاملة، أي كل قيمة ظهرت في التجربة.

ب- طريقة تربيع القيم:

$$SSO = \sum_{i=1}^{t}\sum_{j=1}^{r} Xij^2 - CF \quad \text{أو} \quad SSO = \sum_{i=1}^{n} Xi^2 - CF$$

٣) بعدها نقوم بحساب sst على الشكل التالي:

أ-الطريقة المباشرة:

$$SST = r\sum_{i=1}^{t}(\overline{X}i. - \overline{X}..)^2$$

$\overline{X}i$: متوسط المعاملة

ب-طريقة تربيع القيم:

$$SST = \frac{1}{r}\sum_{i=1}^{t} xi.^2 - CF$$

$$SST = \sum_{i=1}^{t}\frac{Ti^2}{r} - CF$$

Ti: مجموع المعاملة

٤) وأخيرا يتم حساب sse على الشكل التالي:

$$Sse = sso - sst$$

بعد حساب sso و sst و sse نقوم بتقسيم درجات الحرية الكلية أيضا طبقا للمصادر نفسها، وذلك على الشكل التالي:

٥- درجات الحرية الكلية المقابلة لsso هي: tr-1 أو N-1

٦- درجات الحرية الكلية المقابلة لsst هي: t-1 أو t-1

٧- درجات الحرية الكلية المقابلة لsse هي: t(r-1) أو N-t

بعد ذلك نقوم بحساب متوسط مربعات الانحرافات، وذلك على الشكل التالي:

٨-متوسط مربعات الانحرافات بين المعاملات:

$$S^2t = \frac{SSt}{t-1}$$

٩-متوسط مربعات الانحرافات داخل المعاملات أو الخطأ التجريبي:

$$S^2e = \frac{SSe}{t(r-1)}$$

وأخيرا نصل إلى قيمة F وهي:

$$F = \frac{S^2T}{S^2e}$$

وتوضع كل المؤشرات الإحصائية السابقة في جدول يسمى جدول تحليل التباين ويسمى جدول ANOVA (Analysis of variance table) كالتالي:

مصدر الاختلاف Source of variation	درجات الحرية Degrees of freedom	مجموع المربعات Sum of squares	متوسطات المربعات Mean of squares	قيمة F Value F
بين المجموعات أو المعاملات Between treatment	t-1	SST	S^2T	$\frac{S^2T}{S^2e}$
داخل المجموعات Within treatment	t (r-1)	SSe	S^2e	
المجموع Total	tr-1	SSO		

وعادة يكتفى بنشر هذا الجدول في المقالات والمجلات العلمية، بدلا من نشر كل أرقام التجربة.

وأخيرا تقارن F المحسوبة مع F الجدولية، وذلك عند درجات الحرية للمتغير المستقل أو بين المعاملات أفقيا، ودرجات حرية الخطأ التجريبي عموديا، وذلك عند احتمال 5% ، واحتمال 1% فإذا كانت:

F المحسوبة > F الجدولية دل ذلك على رفض فرضية العدم، مما يدل على أن هناك فروق معنوية بين المعاملات المدروسة، أي يوجد تأثير للمتغير المستقل، أو وجود علاقة بين هذا المتغير والاستجابة

F المحسوبة < F الجدولية، دل ذلك على قبول فرضية العدم، مما يدل على أنه ليس هناك فروق معنوية بين المعاملات المدروسة.

مثال: في إحدى تجارب القمح زرعت ثلاثة أصناف من البذار، كل في ستة قطع متساوية، وعوملت جميعها معاملة واحدة من حيث الخدمات الزراعية.

المطلوب: معرفة هل الفروق بين إنتاج هذه الأصناف فروق حقيقية، أم راجعة إلى الصدفة، علما بأنه استخدم في هذه التجربة "التصميم العشوائي الكامل" ،وبيانات التجربة كالتالي:

متوسط الصنف	مجموع الصنف	محصول القطع (كغ)						الصنف
$11.5\bar{x}_1$	$68x_1$	11	16	7	12	14	8	A
$9.5\bar{x}_2$	$58x_2$	9	11	8	10	11	9	B
$15\bar{x}_3$	$90x_3$	13	18	12	14	17	16	C
$x = 12$ المتوسط العام	$x = 216$ المجموع الكلي							

الحل:

١- فرض العدم هو أن الفروق بين متوسطات المعاملات التي توجد في التجربة ما هي إلا فروق عشوائية وناتجة عن الصدفة، ولا يوجد فروق حقيقية بينها، أي:

$HO:\overline{x1.} = \overline{x2.} = \overline{x3.}$

٢-حساب معامل التصحيح CF:

$$CF:= \frac{x^2}{tr} = \frac{(216)^2}{3\times 6} = 2592$$

٣-حساب مجموع مربعات الانحرافات الكلية sso:

$$SSO = \sum_{i=1}^{t}\sum_{j=1}^{r} Xij^2 - CF = [X11^2 + X12^2 + \cdots + X36^2] - CF$$

$$= [8^2 + 14^2 + \cdots + 13^2] - 2592 = 184$$

٤-حساب مجموع مربعات الانحرافات بين المعاملات المدروسة sst:

$$SST = \frac{1}{r}\sum_{i=1}^{t} xi^2 - CF = \left[\frac{x1^2 + x2^2 + x3^2}{r}\right] - CF$$

$$= \left[\frac{68^2 + 58^2 + 90^2}{6}\right] - 2592 = 89.33$$

$$= \left[\frac{4624 + 3364 + 8100}{6}\right] - 2592 = 89.33$$

٥-حساب مجموع مربعات الانحرافات داخل المعاملات sse:

Sse=sso-sst=184-89.33=94.67

٦- حساب درجات الحرية d.f :

أ- درجات الحرية الكلية

tr-1=18-1=17

ب- درجات الحرية لمصدر الاختلاف بين المعاملات t-1=3-1=2

ج- درجات الحرية لمصدر الاختلاف داخل المعاملات

(t r-1) - (t-1) = tr-1-t+1=wt r-t=t(r-1)

= 3(6-1) = 15

٧- بعد استخراج المؤشرات الإحصائية السابقة ، توضع في جدول تحليل التباين ويختصر عادة إلى : Anova table

S.O.V	DF	S.O.S	M.S.O.S	F
B.T	2	89.33	$\frac{89.33}{2} = 44.67$	
W.T	15	94.67	$\frac{94.67}{15} = 6.31$	$\frac{44.67}{6.31} = 7.08$
Total	17	184		

وأخيرا نقارن مع F الجدولية وذلك عند درجات حرية t-1 أفقيا و t(r-1) عموديا، وذلك عند احتمال 5% واحتمال 1%

فالمقارنة تتم عند درجات حرية (15,2) ، وبالنظر في الجدول آخر الكتاب نجد أن قيمة F الجدولية على الشكل التالي:

3.68 على مستوى معنوية 5%

6.36 على مستوى معنوية 1%

وطالما أن (F) المحسوبة أكبر من (F) الجدولية على المستويين للمعنوية، فنقول أنه يوجد فرق معنوي جدا بين الأصناف من ناحية المحصول، وأن هذا الفرق حقيقي و مؤكد إحصائيا، وليس للصدفة أي دور في ذلك الأمر الذي نرفض فيه فرض العدم، وبذلك نكون على مدى من الصحة في هذا الاستنتاج في 99% من الحالات.

جدول (٧) توزيع F بمستوى معنوية 5%

F-Distribution $(F_{.95})$

The numbers given in this table are the values of F for which the area to the left equals 0.95 for Table Va, 0.975 for Table Vb, and 0.99 for Table Vc for the indicated numerator and denominator degrees of freedom.

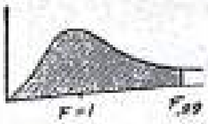
Degrees of freedom for denominator	Degrees of freedom for numerator 1	2	3	4	5	6	7	8	9	10
1	161	200	216	225	230	234	237	239	241	242
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

تابع جدول (٣)

TABLE V . *F-Distribution* $(F_{.95})$
(*continued*)

Degrees of freedom for denominator	Degrees of freedom for numerator								
	12	15	20	24	30	40	60	120	∞
1	244	246	248	249	250	251	252	253	254
2	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.25	2.18	2.10	[illegible].05	2.01	1.96	1.92	1.87	1.81
22	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

جدول (VI) توزيع F بمستوى معنوية 1%



F-Distribution ($F_{.99}$)

	Degrees of freedom for numerator									
Degrees of freedom for denominator	1	2	3	4	5	6	7	8	9	10
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
21	8.02	5.78	4.87	[illegible]	4.04	3.81	3.64	3.51	3.40	3.31
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

تابع جدول (VI)

TABLE Vc. **F-Distribution** ($F_{.01}$)
(continued)

Degrees of freedom for denominator	Degrees of freedom for numerator								
	12	15	20	24	30	40	60	120	∞
1	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Table V is reproduced with the permission of Professor E. S. Pearson from M. Merrington, C. M. Thompson, "Tables of percentage points of the inverted beta (F) distribution," Biometrika, Vol. 33 (1943), p. 73.

١

الفصل التاسع

اعداد : د. قصي العمر

التجارب ذات العامل الواحد Single Factor Experiments

يعتبر هذا النوع من التجارب من أبسط طرز التجارب ،حيث يتم دراسة تأثير عامل متغير واحد على المادة التجريبية المستخدمة (أثر التسميد او الري أو موعد الزراعة أو كمية البذار على محصول معين ، تأثير عدد من العلائق على حيوان معين ، تأثير عدد من المبيدات على حشرة معينة :......الخ) ،ويطلق على هذه التجارب اسم التجارب البسيطة ،وعادة يستخدم لتنفيذ هذا النوع من التجارب التصاميم العشوائية البسيطة التي يمكن بواسطتها دراسة متغير واحد فقط والتصاميم البسيطة هي :

التصميم العشوائي الكامل The completely Randomized design (CRD)

تصميم القطاعات العشوائية الكاملة R andomized complete Blocks design (RCBD)

تصميم المربع اللاتيني Latin square design (LSD)

التصميم العشوائي الكامل CRD

يعتبر التصميم التام التعشية من أبسط التصميمات وأسهلها تحليلا" ، ويستخدم غالبا" عندما تكون الوحدات التجريبية متجانسة (Homogeneous)، أي أن الاختلافات التي بينها تكون ضئيلة ويتوفر هذا التجانس عادة في التجارب المعملية ،ولكن قليلا" ما يتوفر في التجارب الحقلية ،ومن أمثلة التجارب التي يستخدم فيها هذا التصميم ما يلي :

تجارب الأطباق البترية – تجارب الأصص في البيوت البلاستيكية – تجارب تجري على حيوانات من عمر واحد أو وزن واحد.............الخ.

ويعتبر هذا التصميم من أبسط صور التقسيم التي تستخدم في تصميم التجارب ،حيث تقسم فيه مسببات الانحرافات أو الاختلافات أو المجموع الكلي لمربعات الانحرافات SSo الى مصدرين رئيسين هما :

١-الاختلافات التي ترجع إلى مجموعات أو معاملات التجربة ، ويطلق عليها الاختلافات بين المعاملات SST

٢-الاختلافات التي توجد بين أفراد التجربة في داخل المعاملات ،ويطلق عليها الاختلافات داخل المعاملات ،أو الخطأ التجريبي ،أو خطأ العينات ، أو البواقي SSe

أما بالنسبة لكيفية إجراء هذا التصميم ، فيتم ترتيب الوحدات التجريبية في مكان التجربة (حقل – معمل-حجرة –حظيرة حيوانات) عشوائيا" بدون نظام محدد ، سوى أن تعطى لكل وحدة تجريبية نفس الاحتمال في استلام أو تمثيل أية معاملة في التجربة.

ويتم التوزيع العشوائي للوحدات التجريبية على المعاملات إما باستعمال القرعة ، أو باستعمال جداول الأرقام العشوائية.

التعشية باستخدام القرعة :

في تجربة لدراسة أثر خمسة تركيزات من الملوحة (A,B,C,D,E) على صنف الشعير الأسود تم اجراء هذه التجربة بواقع ثلاث مكررات لكل معاملة مثلا" ، وبالتالي فان التوزيع العشوائي يكون كالتالي :

١-حساب عدد الوحدات التجريبية التي تطبق في هذه التجربة وهي :

5 معاملات × 3 مكررات = 15 وحدة تجريبية .

٢-يتم كتابة أرقام الوحدات التجريبية من /1 - 15/ على قطع متساوية من الورق ، ثم تطوى ، وتوضع في سلة أو صندوق،وتقلب جيدا" .

٣-تسحب الواحدة بعد الأخرى بدون النظر الى السلة ،وتدون الأرقام أولا" بأول.

٤-توزع الوحدات التجريبية على المعاملات ، بحيث تكون الأرقام الثلاثة الأولى للمعاملة A، والثلاثة الثانية للمعاملة B وهكذا ، وذلك على الشكل التالي :

15، 11، 2 - 8 ،13 ،3 – 10، 9،7 – 14 ، 5 ، 4 – 6 ، 12 ،1

E D C B A

•يكون الترتيب للوحدات التجريبية في مكان التجربة على الشكل التالي :

D	E	A
A	B	B
C	D	C
A	E	C
E	B	D

التعشية باستخدام جداول الأرقام العشوائية :

تعتبر من أكثر الطرق استخداماً" لتوزيع المعاملات عشوائياً" على الوحدات التجريبية ، و تتلخص هذه الطريقة فيما يلي :

١-اختيار نقطة بداية في جداول الارقام العشوائية (الجداول موجودة آخر الفصل)

٢-التحرك في أي اتجاه للأرقام (يمين – يسار – أعلى – اسفل) ، مع تسجيل الأرقام الثنائية الخمسة عشر الأولى ، وفي حالة تكرار ظهور أي رقم نأخذ الرقم الثنائي الذي يليه .

٣- ترتيب الأرقام العشوائية تصاعدياً" ، مع تسجيل الرتبة لكل رقم في عمود جديد، ويمثل هذا العمود عملية تعشية للأرقام الخمسة عشر من 1 الى 15 التي تمثل الوحدات التجريبية .

٤-تخصيص الوحدات التجريبية الثلاثة الأولى للمعاملة A، ثم الوحدات الثلاثة التالية للمعاملة B، ثم C، ثم D ، ثم تخصص الوحدات التجريبية الثلاثة الأخيرة للمعاملة E.

فلو أخذنا نقطة البداية عند العمود V وفي منتصفه عند الرقم 5 ، وبذلك تتلقى الوحدات التجريبية المرقمة 2 و 12 و 9 المعاملة A ، والوحدات 6 و 13 و 11 المعاملة B، والوحدات 7 و 14 و10 المعاملة C ، والوحدات 4 و 15 و3 المعاملة D، وأخيراً" الوحدات 1 و 8 و 5 تستلم المعاملة E ، كما هو واضح في الجدول التالي :

المعاملة Treatment	الرتبة Rank	الترتيب order	الرقم العشوائي Random numbers	الرقم number
A	2	4	05	1
A	12	5	46	2
A	9	7	35	3
B	6	11	17	4
B	13	12	47	5
B	11	17	40	6
C	7	30	30	7
C	14	32	55	8
C	10	35	36	9
D	4	36	11	10
D	15	40	58	11
D	3	46	07	12
E	1	47	04	13
E	8	55	32	14
E	5	58	12	15

ويوضح الجدول التالي تخطيط تجربة في التصميم العشوائي الكامل بخمس معاملات وثلاثة تكرارات (نفس المثال السابق) .

D	A	E	
B	E	D	
A	E	C	
A	B	C	
D	C	B	

استعمالات هذا التصميم :

يستعمل هذا التصميم في الحالات التالية :

١-التجارب التي تكون فيها وحدات التجربة متجانسة تماما" (حقل متجانس الخصوبة – حيوانات ذات وزن أو عمر واحد) ، بحيث يؤدي ذلك الى تقليل قيمة الخطأ التجريبي ، لأن الفروق بين هذه الوحدات التجريبية المتجانسة سوف تكون صغيرة .

٢-التجارب التي يتعذر فيها على المجرب ترتيب الوحدات التجريبية في مجاميع (وحدات تجريبية ذات خصوبة جيدة – ذات خصوبة متوسطة – ذات خصوبة منخفضة) كما في أنواع التصميمات الأخري ، لعدم توفر أي بيانات أو معلومات تساعد على عمل هذا الترتيب .

٣-التجارب التي يقل فيها المجموع الكلي لدرجات الحرية عن ٢٠ ،ففي هذه التجارب الصغيرة فان نقص عدد درجات الحرية المقابلة للخطأ التجريبي يؤدي إلى كبر قيمة F الجدولية اللازمة لاختبار المعنوية ، وبذلك لاتظهر معنوية الفروق الصغيرة نسبيا" ، وهذا يقلل من حساسية التجربة ، وهذا كله في حال استعمال تصميم آخر غير هذا التصميم .

٤-التجارب التي يحتمل فيها عدد كبير من القيم الغائبة (موت بعض النباتات أو حيوانات التجربة) .

وكما سبق وذكرنا فان هذا التصميم شائع الاستعمال في تجارب الأطباق والأصص والمعامل البحثية والبيوت البلاستيكية ، والتجارب التي تجري على بعض الحيوانات شريطة وضع هذه الحيوانات في حظيرة واحدة ، ولكن هذا التصميم لايستعمل كثيرا" في التجارب الحقلية ، لأن الحقل المتجانس تماما" نادر الوجود ، حيث وجد أن كثيرا" من أراضي

محطات التجارب في العالم غير ملائمة للتجارب الحقلية ، وذلك للتباين الكبير في الأرض الا انه يكفي أن توجد بعض الأجزاء المتجانسة بالمحطة البحثية لتخصيصها للتجارب .

مزايا التصميم :

١-يسمح باستعمال اي عدد من المعاملات باي عدد من المكررات .

٢-سهولة التحليل الإحصائي للتجربة ، حتى بفرض عدم تساوي مكررات المعاملة، أو غياب بيانات بعض الوحدات التجريبية للمعاملة ، أو حتى المعاملة بأكملها ، لأن التأثير النسبي لغياب هذه البيانات يقل من اي من التصميمات الأخرى .

عيوب التصميم :

١-العيب الذي يؤخذ على هذا التصميم مبني على أساس دقة التجربة ، إذا أن التوزيع العشوائي الكامل للمعاملات على الوحدات التجريبية ، وتقسيم مصادر الاختلاف إلى مصدرين فقط ، يجعل قيمة الخطأ التجريبي كبيرة ، الا انه يجب أن نأخذ في الأعتبار أن عدد درجات الحرية المقابلة للخطأ التجريبي في هذا التصميم تكون أكبر منها بالنسبة لأي تصميم أخر ، وبالتالي يزيد من حساسية التجربة Accuracyوالتي تزيد بزيادة عدد درجات الحرية المقابلة للخطأ التجريبي .

٢-انخفاض كفاءة التصميم في حالة عدم تجانس الوحدات التجريبية .

التحليل الإحصائي لبيانات التصميم التام التعشية في حالة تساوي عدد المكررات :

في تجربة لمقارنة ٤ علائق للتسمين ، أحضر الباحث ٤٠ حيوانا" متجانسا" ، قسمت عشوائيا" الى أربع مجموعات بعد جمعها في حظيرة واحدة ، حيث تم تقديم عليقة معينة لكل مجموعة ، ثم قدرت الزيادة في الوزن بالكغ بعد انتهاء فترة التجربة ، والنتائج المتحصل عليها من التجربة كانت كالتالي :

متوسط المعاملة	مجموع المعاملة	الزيادة في الوزن (كغ)										المعاملات
4	40	6	4	2	4	5	1	2	5	5	6	A
7	70	3	7	8	11	7	4	8	6	9	7	B
10	100	10	11	8	10	11	11	12	9	12	6	C
5	50	7	6	7	7	2	2	3	5	4	7	D
$\bar{X}_{..}=6.5$	$X_{..}=260$											

والمطلوب : اختبار معنوية الفروق بين المعاملات الأربع ؟

يلزم لتحليل النتائج ، عمل جدول تحليل التباين وحساب القيم التالية :

١-حساب معامل التصحيح C.F

$$C.F = \frac{G^2}{N} = \frac{X_{..}^2}{tr} = \frac{260^2}{4\times10} = \frac{67600}{40} = 1690$$

٢-حساب مجموع مربعات الانحرافات الكلية SSo: وهنا توجد لعمليات الحساب طريقتان أحدهما يطلق عليها الطريقة المباشرة $Sso=\sum(x-\bar{x})^2$ ، والاخرى طريقة التربيع القيم ، وكلاهما تصل الى نتائج متساوية ، فلو استخدمنا طريقة تربيع القيم نجد أن :

$$SSo=\sum Xi^2 - c.f=[X1^2+X2^2+.........+Xi^2]-c.f=$$

$$=[6^2+5^2+ \ldots\ldots 6^2+7^2] - 1690=2048-1690=358.$$

٣-حساب مجموع المربعات الانحرافات بين المعاملات المدروسة SST :وهنا نلاحظ أن لدينا أربع معاملات (أربع علائق) ، وعليه فان الحساب قد يكون بالطرقة المباشرة

$$SST=r\sum_{i=1}^{t}(\bar{x}-\bar{\bar{x}}\ \ )$$

$$= 10[(4-6.5)^2 + (7-6.5)^2 + (10-6.5)^2 + (5-6.5)^2] =$$

$$=10(6.25+0.25+12.25+2.25)=10\times21=210$$

أو بطريقة تربيع القيم بحيث أن :

$$SST=\sum\frac{Ti^2}{ri} - c.f= \ \ \frac{T1^2+T2^2+T3^2+T4^2}{r} - \frac{G^2}{tr} =$$

$$=\frac{40^2+70^2+100^2+50^2}{10} - \frac{260^2}{4\times10} = \frac{19000}{10} - 1690 = 210$$

٤-حساب مجموع مربعات الانحرافات داخل المعاملات SSe وهي مجموع مربعات الانحرافات للخطأ التجريبي

$$SSe= SSo - SSt = 358 - 210 = 148$$

٥-حساب درجات الحرية d.f :

أ-درجات الحرية لمصدر الاختلاف رقم (١) هو عدد كل أفراد التجربة ناقصا" واحد

$$t.r -1 = 40 - 1 = 39$$

ب-درجات الحرية لمصدر الاختلاف رقم (٢) هو عدد المعاملات ناقصا" واحد

$$t - 1 =4 - 1 = 3$$

ج-درجات الحرية لمصدر الاختلاف رقم (٣) وهو :

$$t(r-1) = 4(10 - 1) = 4\times 9= 36$$

بعد استخراج المؤشرات الإحصائية السابقة ، توضع في جدول ANOVA table

مصدر التباين S.O.V	درجات الحرية d.f	مجموع مربعات الانحرافات S.O.S	متوسط مربعات الانحرافات M.S.O.S	ف F
بين المعاملات Between	3	210	70	$F=\frac{st^2}{se^2}$ $=\frac{70}{4.11}$ = 17.03
داخل المعاملات (الخطأ التجريبي) Within	36	148	4.11	
المجموع Total	39	358		

ثم نقارن F المحسوبة مع جدول F النظرية ، وذلك عند درجات حرية t-1 أفقيا" و t(r-1) عموديا" وذلك عند احتمال5% واحتمال1% وبالمقارنة نجد أن F الجدولية عند مستوى 1%=4.41 وعند مستوى 5%=2.88 اي أن F المحسوبة أكبر من F النظرية على المستويين وهذا يدل على رفض النظرية الفرضية ، مما يدل على أن هناك فروق معنوية [illegible] بين المعاملات الأربع المدروسة ، أي نرفض فرض العدم ونقبل الفرض البديل ،

التحليل الإحصائي لبيانات لتصميم التام التعشية في حالة عدم تساوي التكرارات :

يستخدم هذا التحليل في التجارب الحقلية التي يتوقع فيها تلف أو موت بعض الوحدات التجريبية اثناء اجراء التجربة .

وجدول تحليل التباين في حالة عدم تساوي التكرارات هو نفس الجدول لحالة تساوي التكرارات مع بعض التعديلات الطفيفة وهي :

-١

$$c.f = \frac{X..^2}{r..}$$

$r..$:العدد الكلي لقيم التجربة .

-٢

$$SST = \sum_{i=1}^{t} \frac{Xi.^2}{ri.} - c.f$$

$ri.$:عدد مرات تكرار المعاملة

٣- درجات الحرية الكلية المقابلة ل SSo :

$$r.. - 1$$

٤- درجات الحرية المقابلة للخطا التجريبي SSe :

$$r..-t$$

المقارنات الفردية Individual coparisons :

بعد الانتهاء من إجراء تحليل التباين لنتائج أي تجربة ، يواجه الباحث بإحدى الحالتين :

أ-أن تكون الفروق بين متوسطات المعاملات غير معنوية أي f المحسوبة < f الجدولية وفي هذه الحالة نتوقف عن إجراء أي اختبارات.

ب-أن تكون الفروق بين متوسطات المعاملات معنوية أي f المحسوبة > f الجدولية وهذا يتطلب منا مقارنة متوسطات المعاملات لنتمكن من مقارنة المعاملات والتعرف على أحسنها وهناك العديد من الطرق المستخدمة لإجراء المقارنات الفردية بين متوسطات المعاملات وتختلف في ظروف استعمالها وكذلك في دقتها وسوف نستعرض بعض هذه الطرق مع الإشارة إلى أن هذه المقارنات مشروطة بمعنوية f في جدول تحليل التباين .

اختبار LSD أو أقل فرق معنوي : The Least significant difference

اكتشف العالم Fisher هذا الاختبار لمقارنة الفروق بين مجموعة من متوسطات المعاملات بمتوسط معاملة الشاهد أو المعاملة القياسية وهذا سوف يمكننا من معرفة معنوية الفروق بين المعاملات وأفضل المعاملات فمثلا" مقارنة متوسطات أصناف جديدة بمتوسط الصنف المحلي السائد في المنطقة .

والطريقة التي يستخرج بها قيمة أقل فرق معنوي يتطلب منا الاستعانة بمعادلة t وهي :

$$T = \frac{\overline{x1.} - \overline{x2.}}{\overline{sd}} \Longrightarrow \overline{x1.} - \overline{x2.} = t . \overline{sd}$$

ويتم التعويض عن $\overline{x1.} - \overline{x2.}$ بالأحرف LSD أي أقل فرق معنوي والذي من خلاله يتم مقارنة متوسطات المعاملات : ويتم حساب مؤشرات هذه المعادلة كما يلي :

١حساب t : يتم حساب t من خلال جداول التوزيع الخاصة بها وعند درجات حرية مقابلة للخطأ التجريبي أي عند $t(r-1)$
حيث يتم اخذ قيمة t على مستويي المعنوية 5% ،1% وبالعودة الى المثال السابق والرجوع الى جداول t عند درجات الحرية 36 نجدها :

$$t0.05_{36} = 2.031 \qquad , \qquad t0.01_{36} = 2.727$$

٢حساب $\overline{sd}$:وهو الخطأ القياسي (المعياري) للفرق بين متوسطين ، ويرمز له بـ$\overline{sd}$ تمييزا" له عن الخطأ المعياري المتوسط الحسابي العادي $\overline{sx}$ ويحسب كما يلي :

ا-في حالة تساوي عدد أفراد المعاملتين قيد المقارنة او الوحدات التجريبية (المكررات) للمعاملتين أي r1=r2 :

$$\overline{sd} = \sqrt{\frac{2se^2}{r}}$$

se^2 : مربع الانحرافات القياسي أو التباين للفرق بين متوسطين أو متوسطات مربعات الخطأ التجريبي .

r : عدد الافراد في المعاملة او عدد المكررات في التجربة .

ب-في حالة عدم تساوي أفراد المعاملتين قيد المقارنة أو الوحدات التجريبية للمعاملتين أي r1≠r2 :

$$\overline{sd} = \sqrt{se^2\left(\frac{1}{r1}+\frac{1}{r2}\right)}$$

r1 : عدد الافراد في المعاملة الاولى .

r2 : عدد الافراد في المعاملة الثانية .

وبالرجوع الى المثال السابق نجد أن :

$$\overline{SD} = \sqrt{\frac{2\times4.11}{10}} = \sqrt{0.822} = 0.907$$

٣- حساب LSD عند مستويين المعنوية :

$$LSD_{0.05} = t_{0.05} \cdot \overline{sd} = 0.907\times 2.031 = 1.842$$

$$LSD_{0.01} = t_{0.01} \cdot \overline{sd} = 0.907\times 2.727 = 2.473$$

٤-حساب المتوسطات الحسابية للمعاملات :

$$\bar{x} = \frac{\sum x}{r} = \frac{\text{مجموع قيم المعاملة}}{\text{عدد تكرار المعاملة}}$$

وبالعودة الى المثال نجد أن :

$\bar{A}=4$ ، $\bar{B}=7$ ، $\bar{c}=10$ ، $\bar{D}=5$

٥-اجراء المقارنات :

حيث نقارن الفرق بين متوسطات المعاملات بقيمة أقل فرق معنوي ولمستويي المعنوية 5%و 1%

وبذلك نحكم على أحسن معاملة مع مراعاة ما يلي :

١-المقارنة تتم بالقيمة المطلقة والعلامات تدل على التالي :

** وجود فروق معنوية عالية

* وجود فروق معنوية

- عدم وجود فروق معنوية

٢-ان استعمال اختبار LSD يكون قاصرا" على الحالات التي تكون فيها المقارنة بين صنف محلي وعدد من الاصناف الجديدة مثلا" او بين متوسط معاملة الشاهد ومتوسط المعاملات الاخرى في التجربة .

٣-ان استعمال هذا الاختبار يحتم عدم ظهور أية معاملة في عملية المقارنة الا لمرة واحدة وبالعودة الى المثال واعتبار المعاملة A هي الشاهد فان عملية المقارنة تجرى كما يلي :

المعاملات Treatments	الفروق بين المتوسطات Differences	LSD	
		0.05 =1.84	0.01=2.47
A – B	4 – 7 = -3	*	**
A – C	4 – 10 = -6	*	**
A – D	4 – 5 =-1	-	-

والنتيجة النهائية :

نرى من خلال المقارنة تفوق المعاملة B على A

وتفوق المعاملة C على A والفروق الباقية ليست ذات احصائية وعليه ووفقا" لهذا التحليل نستطيع ان نجيب على تساؤلنا حول أي معاملة أحسن ويكون الجواب هو :

C ثم B ، A

D

مثال : أجريت تجربة لمقارنة محصول خمسة أصناف من القطن كغ/هـ تحت الظروف الطبيعية واستخدام التصميم التام التعشية (CRD) بخمسة تكرارات ولكن إثناء التجربة تلفت بعض الوحدات التجريبية وتمثل البيانات التالية نتائج هذه التجربة :

المعاملة	A	B	C	D	E	
التكرارات	670	440	220	630	450	
	530	290	210	790	560	
	650	340	255	650	410	
	570	375		625	360	
	690					
$\sum X_{i.}$ r_i	3110 5	1445 4	685 3	2695 4	1780 4	9715 20
$\overline{X_{i.}}$	622	361.25	228.33	673.75	445	485.75

والمطلوب : اختبار معنوية الفروق بين المعاملات الخمسة ؟

الحل :

$$c.f=\frac{X_{..}^2}{r_{..}} = \frac{9715^2}{20} = 4719061.2$$

$$SSo = \sum_{i=1}^{t} \sum_{j=1}^{r} Xij^2 \;-\; cf \;=\; [\,670^2 + 530^2 + \ldots\ldots + 360^2\,] - cf =$$

$$= 5292675 - 4719061.2 = 573613.8$$

$$sst = \sum \frac{xi.^2}{Ri} - CF = [\frac{3110^2}{5} + \frac{3445^2}{4} + \frac{685^2}{3} + \frac{2695^2}{4} + \frac{1780^2}{4}] - CF =$$

$$= [1934420 + 822006.25 + 156408.33 + 1815756.2 + 792100] - CF =$$

$$= 5220690.7 - 4719061.2 \quad = \quad 501629.5$$

$$SSe = SSo - SSt = 71984{,}3$$

$$St^2 = \frac{SSt}{T-1} = 125407.37$$

$$Se^2 = \frac{SSe}{r-t} = 4798.95$$

$$F = \frac{St^2}{Se^2} = 26.13$$

وتلخص هذه الحسابات في جدول تحليل التباين كالتالي :

S.O.V	d.f	s.s	M.S	F
B-T	4	501629.5	125407.37	26.13
W- T	15	71984.3	4798,95	
Total	19			

ان قيمة F الجدولية على درجات حرية أفقيا" 4 وعموديا" 15 هي :

$$f_{0.01} = 4.89 \quad , \quad f_{0.05} = 3.06$$

وبالمقارنة : f المحسوبة > f الجدولية عند مستويي المعنوية 0.01 و 0.05 ونستخلص من هذه التجربة ان هناك اختلافات معنوية جدا" بين متوسطات الاصناف الخمسة .

قارن بين الاصناف المدروسة باستخدام اختبار LSD علما" أن معاملة الشاهد هي المعاملة E ؟

الحل : نظرا" لأن المكررات هي 5و 4و 3 فاننا يجب ان نحسب ثلاث قيم للخطأ القياسي للفرق :

١-اذا كانت الوحدات التجريبية متساوية في عددها أي r1=r2 فان :

$$\overline{sd} = \sqrt{\frac{2Se^2}{r}} = \sqrt{\frac{2\times4798.95}{4}} = 48.98$$

٢-اذا كانت الوحدات التجريبية غير متساوية في عددها r1≠r2 والمكررات 4و 5

$$\overline{sd} = \sqrt{\frac{4798.95}{5} + \frac{4798.95}{4}} = 46.47$$

٣-اذا كانت الوحدات التجريبية غير متساوية في عددها r1≠r2 والمكررات 4و 3:

$\overline{sd} = \sqrt{\frac{4798.95}{3} + \frac{4798.95}{4}} = 52.91$

ثم نحسب LSD للحالات الثلاثة وذلك بضرب الخطأ المعياري $\overline{sd}$ في كل من

$t_{0.01} = 2.947$ ، $t_{0.05} = 2.131$

١-متساوية

$LSD_{0.05} = t_{0.05} \cdot \overline{sd} = 2.131 \times 48.98 = 104.38$

$LSD_{0.01} = t_{0.01} \times \overline{sd} = 2.947 \times 48.98 = 144.34$

٢-المكررات 4 و 5

$LSD_{0.05} = 99.03$

$LSD_{0.01} = 136.95$

٣-المكررات 4 و 3

$LSD_{0.05} = 112.75$

$LSD_{0.01} = 155.93$

ثم نجري مقارنة الاصناف بمعاملة الشاهد وهي E

5-4 متساوية

3-4 متساوية

Treat	Differences	Significance	LSD 0.05	LSD 0.01
A – E	622-445=177	**	99.03	136.95
B – E	361.25-445=83.7$\overline{3}$	-	104.38	144.34
C – E	228.33-445=216.6$\overline{7}$	**	112.75	155.93
D – E	673.75-445=228.75	**	104.38	144.34

من خلال الجدول نلاحظ تفوق A على E

تفوق D على E.

تفوق E على C.

عدم وجود فروق معنوية بين E و B

وعلى ذلك ترتيب المعاملات كالتالي : D ، A ، B ، C

E

جدول (I) الارقام العشوائية

Table 8. Sets of Random-Sample Numbers*
(Numbers 1 to 60 inclusive)

I	II	III	IV	V	VI	VII	VIII	IX	X
39	21	17	01	60	02	43	25	29	39
15	02	40	57	06	57	42	46	48	38
34	42	15	30	37	22	50	08	59	15
11	47	45	33	22	27	22	53	28	10
49	33	26	52	26	05	23	30	15	27
33	08	31	56	01	38	29	50	25	51
16	07	36	03	25	59	56	54	19	34
19	53	01	39	02	54	35	23	04	14
12	44	42	24	09	21	19	40	09	05
56	20	05	37	53	29	36	02	30	58
48	39	38	50	56	11	55	60	22	02
26	41	13	32	08	20	38	44	01	20
32	15	25	58	03	32	39	59	41	49
41	03	10	43	50	58	15	13	11	52
59	31	14	27	27	18	09	42	42	59
60	57	11	29	20	45	02	14	02	26
50	59	06	10	42	31	28	35	23	48
27	55	56	11	31	39	46	17	26	04
28	01	52	41	44	30	03	49	47	37
40	29	34	22	21	23	05	22	37	12
37	35	09	13	29	01	53	06	34	55
36	36	49	34	41	14	04	16	44	33
08	48	60	04	18	48	21	55	33	16
57	12	07	51	10	15	50	07	27	08
22	56	30	08	57	06	12	47	13	53
24	05	59	20	54	52	07	45	39	35
55	10	04	48	43	40	32	31	03	09
31	06	19	44	51	55	08	21	55	23
14	14	53	25	05	25	51	52	57	19
45	16	44	40	46	49	54	27	38	21
05	37	57	42	35	08	49	43	52	13
30	30	58	17	17	26	13	12	24	17
17	43	50	07	47	16	10	38	53	03
44	13	28	36	40	19	26	56	45	57
25	45	43	18	30	41	48	34	49	40
47	22	03	14	55	35	06	32	10	44
58	49	48	35	36	50	24	48	54	24
53	58	08	38	11	13	60	10	16	46
43	32	20	49	58	53	33	41	14	50
21	04	21	23	07	10	47	26	43	30
20	27	51	09	04	43	35	09	17	01
23	11	54	12	32	09	44	29	35	06
42	46	22	47	12	17	16	28	32	07
52	18	18	15	49	36	25	57	56	42
13	23	55	19	28	07	57	01	40	22
54	28	37	59	45	03	30	24	60	60
07	51	23	45	34	44	45	39	50	41
06	24	16	06	48	60	17	37	20	58
02	50	35	55	24	34	18	15	06	31
35	26	47	60	59	47	20	36	18	32
29	17	41	31	14	28	37	05	36	11
04	09	33	05	15	24	52	58	58	45
18	19	39	02	52	33	40	04	46	18
46	60	24	26	33	56	01	51	31	29
10	34	32	21	19	51	31	11	51	54
09	54	27	46	23	12	41	03	05	25
03	25	29	53	38	37	11	20	12	28
01	38	12	28	13	42	58	18	07	47
38	52	02	16	16	47	27	33	21	36
51	40	46	54	39	04	14	19	08	43

إعداد أ.د. قصي العمر

تصميم القطاعات العشوائية الكاملة

Randomized complete blocks design

تطرقنا في الفصل السابق إلى **التصميم العشوائي الكامل** ، ورأينا كيفية توزيع المعاملات على كامل الوحدات التجريبية عشوائيا بدون قيد أو شرط ، حيث من المفروض أن جميع الوحدات التجريبية متشابهة تماما ومتجانسة " **Homogeneous** " ، أي إن هناك شرط أساسي هو **تجانس** الوحدات التجريبية ، **وإذا لم يتوفر هذا الشرط** مثل عدم وجود حقل متجانس أو عدد متشابه من الحيوانات فإن الفروق التي توجد بين الوحدات التجريبية سوف تزيد من قيمة الخطأ التجريبي ، ويؤدي ذلك إلى التقليل من كفاءة التجربة ، وبالتالي إخفاء الفروق الحقيقية بين المعاملات .

لذلك إذا وجدت حالة **عدم التجانس** بين الوحدات التجريبية ، وهي حالة شائعة في العديد من التجارب ، فإننا نلجأ إلى تصميم آخر من شأنه تقسيم مفردات التجربة إلى مجموعات متجانسة من الوحدات التجريبية تسمى كل مجموعة منها **قطاع أو مكررة** ، ثم نقوم بمقارنة المعاملات داخل القطاعات ، وبهذه الطريقة يصبح بإمكاننا استخراج الاختلافات بين القطاعات من الخطأ التجريبي مما يؤدي إلى تصغير الخطأ التجريبي .

ويعتبر هذا التصميم من **أكثر أنواع التصميمات استعمالا في التجارب والأكثر شيوعا في ميادين البحث العلمي** ، لأنه يعطي درجة دقة عالية ، وهو يعتمد على نفس المبدأ الذي ذكرناه في حالة وضع الوحدات التجريبية في أزواج في اختبار(t)

إن الغرض الأساسي من تجميع الوحدات التجريبية المتجانسة في قطاع واحد هو تمكين الباحث من مقارنة المعاملات بدرجة كبيرة ، وغالبا ما تكون هناك إمكانية لتجميع الوحدات التجريبية في قطاعات متجانسة إذا ما توفر لدينا معرفة بعض الظواهر الخاصة بالمادة التجريبية .

وهنالك **العديد من الأفكار** التي على أساسها نقوم بذلك التجميع مثلا بالنسبة للتجارب الزراعية الحقلية تكون الوحدات التجريبية القريبة من بعضها متجانسة من ناحية التربة أكثر من الوحدات البعيدة ، لذلك **تجمع الوحدات القريبة من بعضها في قطاع واحد** ، بحيث إن وجدت اختلافات في الاستجابة (**response**)سيكون مصدرها الاختلافات بين المعاملات وليس الاختلافات التي بين الوحدات التجريبية

وبالنسبة لتجارب الإنتاج الحيواني تجمع الحيوانات حسب الصنف والعمر والجنس او الصفات الوراثية .

والنقطة المهمة في عملية التجميع هي : أن يكون التباين بين الوحدات التجريبية داخل القطاع ، اقل من التباين الذي بين كل الوحدات التجريبية ، وإلا تصبح عملية استخدام القطاعات غير ناجحة ، وإذا لم يتوفر هذا الشرط تكون كفاءة هذا التصميم أقل من كفاءة التصميم التام التعشية .

ونستنتج من المفهوم السابق النقاط التالية :

١) بالنسبة للتجارب الحقلية الزراعية إذا كان هناك ميل في الأرض (الانحدار) او إختلاف في درجه الخصوبة توضع القطاعات بطريقة متعامدة لذلك الاتجاه مع تفضيل القطاعات الطويلة وغير العريضة

مكرر 1	مكرر 2	مكرر 3
A	A	B
B	C	C
C	B	A

ووضع اتجاه المكرر بشكل عمودي على اتجاه الخصوبة أو الانحدار يضمن أن عدم التجانس ينطبق بدرجة متماثلة على كافة القطاعات الخاصة بالتجربة

٢) تفضل القطاعات الصغيرة التي تشمل على اقل عدد من الوحدات التجريبية نظراً لأنه كلما زاد حجم القطاع زاد التباين بين الوحدات التجريبية داخل القطاع ، وهذا غير مستحب لأنه يقلل من دقة التجربة

٣) نحصل على أفضل القطاعات بالتقليل من الاختلافات داخل القطاع وزيادة الاختلافات بين القطاعات ، ولهذا نوحد دائما الإجراءات التنفيذية للقطاع الواحد ومعاملة الوحدات التجريبية للقطاع الواحد بطريقة موحدة ، وإذا ما تمت عملية التجميع توزع المعاملات عشوائيا داخل كل قطاع بطريقة مستقلة عن القطاعات الأخرى في التجربة ، ولذلك فان دقة هذا التصميم تعود إلى قدرته على تقليل قيمة الخطأ التجريبي عن طريق تقسيم مادة التجربة إلى قطاعات يعامل كل منها كأنه تجربة مستقلة . واستناداً إلى ذلك

إعداد : أحمد العلي

إشراف : د. صبحي العبد

يجب أخذ القراءات الحقلية وإجراء جميع العمليات الزراعية على جميع الوحدات التجريبية داخل كل قطاع في نفس الوقت وبنفس الكفاءة كلما أمكن ذلك ، مثلا إذا أردنا استعمال مبيد حشائش معين أو حصاد تجربة معينة وتطلب العمل في هذه التجربة أكثر من يوم ، فإنه ينصح رش جميع القطع التجريبية داخل مكرر أو أكثر في نفس اليوم وكذلك بالنسبة للحصاد ، وعدم ترك بعض الوحدات التجريبية داخل بعض القطاعات بدون معاملة لليوم التالي ، وكذلك الحال عند أخذ النتائج الحقلية حيث يفضل أن يقوم شخص واحد بأخذ القراءات الخاصة بالتجربة أو عدة أشخاص بحيث يتولى كل فرد أخذ القراءات لقطاع واحد أو أكثر .

ميزات وعيوب التصميم:

يستخدم تصميم القطاعات العشوائية الكاملة في حال عدم تجانس الوحدات التجريبية ، حيث يتمكن الباحث من عملية استخراج الاختلافات التي بين القطاعات من الخطأ التجريبي وبالتالي تتحسن دقة التجربة، **والمزايا الأساسية للتصميم هي :**

١) تحسين دقة وكفاءة التجربة باستخدام القطاعات.

٢) إمكانية استخدام أي عدد من المعاملات وأي عدد من القطاعات وتوضع كل معاملة مرة واحدة في كل مكررة أو قطاع .

٣)التحليل الإحصائي لهذا التصميم سهل ومرن إذ يسمح بإلغاء مكررة بأكملها أو بيانات معاملة أو أكثر من المعاملات دون تعقيد لطريقة التحليل وفي حالة غياب قيمة إحدى الوحدات التجريبية(**missing plot**) فانه يمكن تقدير قيمتها حسابياً من بيانات التجربة باستعمال طريقة **(Yates)** إلا أن زيادة عدد الوحدات التجريبية الغائبة تجعل عملية تحليل النتائج أصعب من التصميم العشوائي الكامل.

وهناك بعض العيوب نذكر منها :

١-إن زيادة عدد المعاملات يجعل من الصعب الحصول على قطاع أو مكررة متجانسة ، وهذا يؤدي إلى زيادة قيمة الخطأ التجريبي .

٢-يكون التصميم تام التعشية أفضل من هذا التصميم في حالة تجانس الوحدات التجريبية .

التعشية randomization :

لتوضيح عملية التعشية لتصميم RCBD نفترض أن لدينا T=4 وهي A,B,C,D و r = 3 قطاعات ، فنقسم أولا المادة التجريبية إلي ثلاث قطاعات ويقسم كل قطاع إلى ٤ وحدات تجريبية ثم توزع المعاملات الأربع توزيعاً عشوائياً داخل كل قطاع بطريقة مستقلة عن التعشية التي تقع في القطاعات الأخرى ، والتوزيع يتم إما باستخدام القرعة أو إستخدام جدول الأرقام العشوائية .

و فيما يلي مثال عن تعشية القطاع الأول في التجربة :

لنفرض إننا أخذنا الأرقام الأربعة من جداول الأرقام العشوائية وهي: ٥ ، ٤٦ ، ٣٥ ، ١٧ ثم نقوم بترتيبها تصاعدياً

المعاملة	ترتيب	الرقم العشوائي	رقم
A	1	5	1
B	4	46	2
C	3	35	3
D	2	17	4

ونضع كل معاملة حسب ترتيبها داخل الوحدة التجريبية الخاصة بها ، إي Aفي الوحدة رقم1 وB في الوحدة رقم 4 وC في الوحدة رقم 3 وDفي الوحدة رقم 2

ثم نقوم بنفس العملية للقطاع الثاني والثالث ، وأخيرا يكون شكل التصميم النهائي على النحو الموضح بالشكل التالي ، حيث يشتمل كل قطاع على المعاملات الأربع

ومن هنا أتت التسمية الكاملة للتصميم .

BLOCK I	BLOCK II	BLOCK III
A	C	B
D	A	A
C	D	D
B	B	C

تحليل التباين:

التحليل الإحصائي مبني على أساس تساوي عدد المعاملات داخل كل قطاع أي كل قطاع يحتوي على نفس العدد من الوحدات التجريبية ، وعليه تقسم مجموع مربعات انحرافات المشاهدات إلى ثلاث أجزاء : الأول هو تأثير القطاعات والثاني تأثير المعاملات والثالث الخطأ التجريبي ونلاحظ إن هذا الجدول يشبه **جدول تحليل التباين للتصميم التام التعشية مضافاً إليه مصدر جديد من مصادر الاختلاف وهو** القطاعات

وتحسب مكونات جدول تحليل التباين كالتالي :

١) معامل التصحيح CF :

$$CF = \frac{X^2..}{TR}$$

X.. : المجموع الكلي للقيم الناتجة بالتجربة

R : عدد المكررات

T : عدد المعاملات

٢) مجموع مربعات الانحرافات الكلية SSO :

$$SSO = \sum_{i=1}^{t} \sum_{j=1}^{r} (xij - \overline{x..})^2 \text{ أو } = \sum_{i=1}^{t} \sum_{j=1}^{r} (xij)^2 - CF$$

حيث : $\overline{X..}$ المتوسط العام للتجربة

Xij: كل قيمة ظهرت بالتجربة

٣) مجموع مربعات الانحرافات بين القطاعات **SSR** :

$$SSR = t \sum_{j=1}^{r} (x.j - \overline{x..})^2 = \sum_{j=1}^{r} \frac{(x.j)^2}{t} - CF$$

x.j : مجموع المكرر أو القطاع

$(x.j)^2$: مربع مجموع المكرر

إعداد أحمد العلي

بإشراف د. فضي العمر

٤) مجموع مربعات الانحرافات بين المعاملات **SST** :

$$SST = r\sum_{i=1}^{t}(\bar{x}_{i.} - \bar{x}_{..})^2 = \frac{1}{r}\sum_{i=1}^{t}(x_{i.})^2 - CF$$

$\bar{x}_{i.}$: متوسط المعاملة

$x_{i.}$:مجموع المعاملة

٥) مجموع مربعات الانحرافات داخل المعاملات أو الخطأ التجريبي **SSe:**

$$SSe = \sum_{i=1}^{t}\sum_{j=1}^{r}(X_{ij} - \bar{x}_{.j} - \bar{x}_{i.} + \bar{x}_{..})^2 = SSO - (SSR + SST)$$

٦) درجات الحرية الكلية المقابلة ل SSO **Tr-1**

٧) درجات الحرية الكلية المقابلة ل SSR : **r – 1**

٨) درجات الحرية الكلية المقابلة ل SST : **t – 1**

٩) درجات الحرية الكلية المقابلة ل SSe : **(t – 1)(r – 1)**

ويوضح جدول تحليل التباين كالتالي:

S.O.V	df	SS	MS	F
B-BLOCKS B-Treatment W. Treatment	r – 1 t – 1 (t – 1)(r – 1)	SSR SST SSe	$s_r^2 = \frac{ssR}{r-1}$ $s_t^2 = \frac{SST}{t-1}$ $s_e^2 = \frac{sse}{(t-1)(r-1)}$	$\frac{s_t^2}{s_e^2}$
Total	Tr – 1	SSO		

الأخطاء المعيارية و معامل الاختلاف :

الخطأ المعياري للمتوسط : أو الخطأ القياسي لمتوسط المعاملة.

$$S\bar{x} = \sqrt{\frac{S_e^2}{r}}$$

إعداد/ حمود العلي

بإشراف/ د. هدى العبيد

الخطأ المعياري للفرق بين متوسطين : أو الخطأ القياسي للفرق بين متوسطي معاملتين

$$S\bar{d} = \sqrt{\frac{2S_e^2}{r}}$$

معامل الاختلاف : Coefficient of variation

$$C.V = \left(\frac{\sqrt{S_e^2}}{\bar{X}..}\right)100$$

$\bar{X}..$: المتوسط العام للتجربة

و هو من مقاييس التشتت النسبي و مجرد من وحدة قياس البيانات ، وله مجموعة من الاستخدامات ، ومن أهمها :

١- **يستخدم لمقارنة تجربة بتجربة أخرى من ناحية الدقة** ، فلو كان لدينا تجربتين (تباين الخطأ التجريبي معلوم – والمتوسط العام معلوم) ، فيمكن عن طريق حساب معامل التغير أو الاختلاف تحديد أي التجربتين **أكثر تغيراً** فلو كان :

تجربة٢ $C.V_2 = 7\%$	تجربة ١ $C.V_1 = 17\%$

نستنتج أن التجربة الأولى أكثر تغيراً من التجربة الثانية و بالتالي التجربة الثانية أكثر دقة من الأولى .

٢- **لمعرفة مدى الاعتماد على نتائج التجربة** ، ويفضل في التجارب الحقلية أن لا تزيد قيمته عن ٢٠% بينما في حالة البيوت الزجاجية فهي في حدود ١٠%، وفي المختبرات يفضل أن لا تزيد عن ٥% .

تقدير القيم الغائبة : Estimating missing plots

قد تفقد أحيانا بعض الوحدات التجريبية ، أي المشاهدات لعدة أسباب ، و تختلف هذه الأسباب باختلاف التجارب ، ففي التجارب الحيوانية قد يموت حيوان أثناء التجربة ، وفي التجارب الزراعية قد ترعى حيوانات (أو حيوانات بشرية) نباتات بعض الوحدات التجريبية أو لا تحدث استجابة لبعض الوحدات التجريبية و بذلك

نقص البيانات ، وبما أن التحليل الإحصائي السابق مبني على أساس تساوي عدد المعاملات داخل كل قطاع ، وأن كل قطاع يحتوي على نفس العدد من الوحدات التجريبية ، فيصبح من الضروري تقدير البيانات المفقودة ، و يمكن تقدير قيمة لهذه الوحدة التجريبية من معادلة العالم (Yates) التالية :

ملاحظة: لو فقدت أكثر من وحدة ، يتم وضع قيم تقريبية للوحدات المفقودة ، وتترك واحدة تحسب بطريقة Yates

$$Xij = \frac{r(x_{.j}) + t(x_{i.}) - x_{..}}{(t-1)(r-1)}$$

r : عدد القطاعات

t : عدد المعاملات

$x_{.j}$: مجموع المشاهدات المتبقية داخل القطاع الذي توجد به القيمة المفقودة

$x_{i.}$: مجموع المشاهدات المتبقية من المعاملة التي توجد بها القيمة المفقودة

$x_{..}$: المجموع الكلي للمشاهدات المتبقية

وبعد تقدير المشاهدة المفقودة يتم تحليل البيانات كالمعتاد ، فيما عدا خصم درجة حرية واحدة من كل من درجات الحرية الكلية و درجات الحرية للخطأ التجريبي والجدير بالذكر هنا إن القيمة المقدرة لا تضيف أية معلومة جديدة للبيانات ، ولكن الغرض من هذه العملية هو تمكين الباحث من إجراء تحليل التباين بالطريقة العادية ، ولذلك نوصي بأن لا تدخل القيمة في حساب المتوسطات.

ويقدر الخطأ المعياري للفرق بين متوسط المعاملة ذات القيمة الغائبة ومتوسط أي معاملة أخرى باستخدام المعادلة التالية :

$$S\bar{d} = \sqrt{S^2e\left(\frac{2}{r} + \frac{t}{r(t-1)(r-1)}\right)}$$

التحليل الإحصائي :

يمكن توضيح طريقة تحليل تجربة تصميم قطاعات عشوائية كاملة بالمثال الآتي : أجريت تجربة لمقارنة سبعة هجن من الذرة الصفراء (من B إلى H) ، وأحد الأصناف المحلية (A) في تصميم القطاعات العشوائية الكاملة وكررت زراعة كل من هذا الأصناف أربع مرات ، وكانت القطعة التجريبية عبارة عن ثلاث خطوط طول كل منها سبعة أمتار ، والجدول التالي يبين إنتاج القطعة التجريبية من الذرة الصفراء بالكيلو جرام لكل من الأصناف الثمانية المذكورة .

معاملة مكرر	A	B	C	D	E	F	G	H	X.J	X̄.J
1	⑨	11	10	7	16	11	15	20	99	12.37
2	15	10	12	4	19	15	18	21	114	14.25
3	14	22	21	9	13	23	34	18	154	19.25
4	28	12	25	14	29	18	35	26	187	23.37
XI.	66	55	68	34	77	67	102	85	X..	X̄..
X̄I.	16.5	13.75	17	8.5	19.25	16.75	25.5	21.25	554	17.31

المطلوب :

أولاً : اختبر معنوية الفروق بين المعاملات المدروسة باستخدام اختبار F

ثانياً : احسب الخطأ المعياري للفرق بين متوسطين ، ثم احسب الخطأ المعياري للمتوسط الحسابي .

ثالثاً : احسب معامل الاختلاف وفسر الناتج .

رابعاً :على افتراض إن الوحدة التجريبية الأولى من المكرر الأول والمعاملة الأولى فقدت إثناء إجراء التجربة ،احسب قيمة هذه الوحدة بطريقة العالم Yates .

إعداد أحمد العلي

بإشراف د. فوزي العمر

الحل :

أولاً : لتحليل نتائج هذه التجربة باستخدام اختبار F نتبع الخطوات التالية :

١) حساب معامل التصحيح C.F

$$CF = \frac{554^2}{8*4} = 9591.13$$

٢) حساب مجموع مربعات الانحرافات الكلي للتجربة (sso) :

$SSO = \sum_1^{tr} x^2 i - cF = (9^2 + 11^2 + \dots\dots\dots + 33^2 + 26^2) - 9591.13 = 1792.87$

٣) حساب مجموع مربعات الانحرافات بين المعاملات (SST)

$$SST = \frac{\sum_1^t Ti^2}{r} - c.F =$$

$$\frac{66^2+55^2+68^2+34^2+77^2+67^2+106^2+86^2}{4} - 95913 = 710.87$$

٤) حساب مجموع مربعات الانحرافات بين المكررات أو القطاعات (SSR)

$$SSR = \frac{\sum_1^r ri^2}{t} - c.F = \frac{99^2+114^2+154^2+185^2}{8} - 9591.13 = 594.12$$

٥) حساب مجموع مربعات الانحرافات داخل المعاملات أو الخطأ التجريبي (sse)

SSe=sso-(sst+ssR)=1792.87-(710.87+594.12)=487.88

٦) حساب متوسط مربعات الانحرافات لكل من المعاملات والقطاعات والخطأ التجريبي بقسمه مجموع مربعات الانحرافات لكل مصدر على درجات الحرية المقابلة له .

$$s^2t = \frac{sst}{t-1}$$

$$s^2R = \frac{ssR}{r-1}$$

$$s^2e = \frac{sse}{(t-1)(r-1)}$$

٧) حساب قيمة F بقسمة s_t^2 على s_e^2 ، ثم تقارن مع F الجدولية عند درجات الحرية بين المعاملات أفقيا ، وعموديا عند درجات الحرية المقابلة للخطأ التجريبي ، وذلك عند مستوى المعنوية 0.05 ، 0.01

٨) تلخص النتائج السابقة في جدول تحليل التباين ANOVA Table

S.O.V	d.f	s.s	Ms	F
Between treatment	t-1 7	ssT=710.87	s_T^2 = 101.55	s_T^2
Between Blocks	r-1 3	ssR=594.12	s_R^2 = 198.04	s^2e
Within treatments	(t-1)(r-1) 21	sse=487.88	s^2e = 23.23	4.37
Total	Tr-1 31	Sso=1792,87		

وحيث إن قيمة (F) من الجدول الدرجات حرية 7، 21 هي :

F0.05(7.21)=**2.49**

F0.01(7.21)=**3.64**

إذا يوجد فرق معنوي جدا Highly significant بين المعاملات ، وبناء على ذلك يمكن عمل المقارنات الفردية باستعمال اختبار LSD لمقارنة متوسطات المعاملات بمتوسط معاملة الشاهد وهي المعاملة(A) ، أو باستعمال اختبار دانكان لمقارنة متوسطات جميع المعاملات لتحديد أفضل معاملة .

ثانياً :

$$\text{للمتوسط } S\overline{X} = \sqrt{\frac{S^2e}{r}} = \sqrt{\frac{23.23}{4}} = 2.41$$

$$\text{للفروق بين متوسطين } S\overline{d} = \sqrt{\frac{2S^2e}{r}} = \sqrt{\frac{2*23.23}{4}} = 3.41$$

ثالثاً :

$$\text{C.V} = \left(\frac{\sqrt{S_e^2}}{\bar{X}_{..}}\right)100 = \left(\frac{\sqrt{23.23}}{17.31}\right)100 = 27.84\%$$ معامل الاختلاف

هذا يعني إن النتائج لا يمكن الاعتماد عليها أو الوثوق بها لأنها تزيد عن ٢٠ %

رابعاً:

تقدير القيمة المفقودة : إذا فرضنا وكانت القيمة الأولى في التجربة السابقة (٩) غائبة فهي تخص المعاملة(A) و المكرر الأول ، فإننا نقدرها بالطريقة التالية :

$$X_{ij} = \frac{4(90) + 8(57) - 545}{(4-1).(8-1)} = 12.9$$

وتستعمل هذه القيمة في تحليل نتائج التجربة كالمعتاد فيما عدا طرح درجة حرية واحدة من درجات الحرية الكلية و الخطأ التجريبي

$[(t-1)(r-1)] - 1\ ,\ \ (tr-1) - 1]$، و ذلك لان متوسط مربعات الانحرافات للمعاملات S_T^2 تميل لأن تكون قيمته مرتفعة، وللتغلب على ذلك تزداد قيمة الخطأ التجريبي بتقليل عدد درجات حريته درجة واحدة .

اختبار دانكان:Duncan test

أوجد دانكان هذه الطريقة عام ١٩٥٥م ، في طريقة مشابهة لطريقة أقل فرق معنوي ، إلا أنها أدق منها خصوصاً إذا أريد عمل كل المقارنات الفردية بين متوسطات المعاملات ومقارنة أكبرها بأصغرها ، وبالذات إذا زاد **العدد عن ٥ معاملات** ، إلا إن هذا الاختبار لا يستعمل غالباً إلا في حالة تساوي عدد المكررات (الوحدات التجريبية) في المعاملات ، وفي حالة عدم التساوي فانه يستخدم الوسط التوافقي $\bar{X}h$ (Harmonic mean) بدلاً من r في حساب $s\bar{x}$ ، والاساس الذي تبنى عليه هذه الطريقة هو مقارنة المدى لمجموعة من المتوسطات.

$$\text{L.S.R} = \sqrt{2}\ .\ t_{(v_2)}\ .\ s\bar{x}$$

$$\text{L.S.R} = \sqrt{2}\ .\ t_{(v_2)} \cdot \sqrt{\frac{s^2e}{r}}$$

s^2e :تباين الخطأ التجريبي

r : عدد الوحدات التجريبية للمعاملة (عدد المكررات)

إعداد محمد العلي

إشراف صبحي العمر

وفي حال طريقة دانكان حيث يقارن أكثر من متوسطي معاملتين ، فإن القيمة التي تقابل أقل فرق معنوي يطلق عليها **Least significant range** وتختصر **L.S.R** ، وقد حسبت قيم $(\sqrt{2} . t)$في معادلة أقل فرق معنوي السابقة عند مستوى المعنوية **0.05** , **0.01** وأطلق عليها دانكان **Significant studentized range** وتختصر **(SSR)** ، ووضعت في جداول خاصة بحيث تمثل القيمة (P) في أعلى الجدول عدد المتوسطات التي تقارن في مجموعة واحدة ، وأمام درجات الحرية للخطأ التجريبي

ولإيجاد قيمة L.S.R نضرب قيمة SSR في قيمة الخطأ المعياري $s\bar{x}$

$$LSR_{0.05} = SSR_{0.05} \cdot s\bar{x}$$

$$LSR_{0.01} = SSR_{0.01} \cdot s\bar{x}$$

بالعودة إلى مثال ٤ علائق للتسمين A,B,C,D ، ٤٠ حيوان ، كل معاملة مكررة ١٠ مرات ، $S^2e = 4.11$ ، درجات حرية الخطأ التجريبي = **36**

١- **حساب SSR** : تستخرج قيمة المعامل SSR من جداول دانكان و عند مستوى المعنوية المطلوبة ، حيث قيمة P من 2 إلى عدد المعاملات المستخدمة في التجربة ، وحسب درجات الحرية للخطأ التجريبي

ففي مثالنا السابق نستخرج قيمة المعامل SSR من جداول دانكان ، عند درجات الحرية الخاصة بالخطأ التجريبي (36) ، ولأعداد من المتوسطات هي 2,3 ، وعند مستوى المعنوية المطلوب

P	2	3	4
SSR_{05}	2.87	3.02	3.11
SSR_{01}	3.86	4.02	4.13

٢- حساب $s\bar{x}$: حيث يحسب الخطأ القياسي حسب المعادلة :

$$s\bar{x} = \sqrt{\frac{S^2e}{r}} = \sqrt{\frac{4.11}{10}} = 0.64$$

- **في اختيار L.S.D نستخدم $S\bar{d}$ أي الخطأ المعياري للفرق بين متوسطين ، لان المقارنة تتم على أساس الفرق بين المتوسطين**
- **في اختيار L.S.R نستخدم $S\bar{X}$ أي الخطأ المعياري للمتوسط الحسابي ، لان المقارنة تتم على أساس عدد المتوسطات المحصورة ضمن مدى المقارنة P**

٣- حساب **L. S. R** : وهو قيم أقل مدى معنوي ، حيث يتم تطبيق المعادلة وحساب قيمة L. S. R لكافة المعاملات ، وعند مستوى المعنوية 0.05 , 0.01 وتوضع في جدول على الشكل التالي :

4	3	2	P
1.99	**1.93**	**1.84**	**LSR_{05}**
2.64	**2.57**	**2.47**	**LSR_{01}**

٤- نحسب متوسطات المعاملات ، ونقوم بترتيبها تصاعدياً كما يلي :

A	**D**	**B**	**C**	**المعاملة treat**
$\frac{40}{10} = 4$	$\frac{50}{10} = 5$	$\frac{70}{10} = 7$	$\frac{100}{10} = 10$	**المتوسط *mean***

٥- **ترتيب المتوسطات مع قيمة L. S. R المناسبة** : بناءً على عدد المتوسطات في مدى المقارنة (P) **ثم إجراء المقارنة بهذه القيم** ، ومن ثم تحديد معنوية الفروق بين المتوسطات كما يلي :

p	المقارنة comparison	الفرق difference	LSR_{05}	LSR_{01}
4	C – A	6**	1.99	2.64
3	C – D	5**	1.93	2.57
2	C - B	3**	1.84	2.47
3	B – A	3**	1.93	2.57
2	B - D	2*	1.84	2.47
2	D - A	1 -	1.84	2.47

ونظراً لضخامة الجداول من هذا النوع خصوصاً في حالة زيادة عدد المعاملات ، فإنه يمكن تلخيص [illegible] في الجدول التالي ، حتى يتسنى نشره في الدوريات العلمية وفي الكتب ، على أن يؤخذ في الاعتبار :

١- يعطى الحرف a للمتوسط الأكبر ثم أحرف الأبجدية الانكليزية (g,f,e,d,c,b , ...) للمتوسطات الأقل .

٢- تعطى المعاملة الأخيرة نفس الحرف الخاص بالمعاملة التي تسبقها إن لم يكن بينهما فرق معنوي ، وان كان هناك فرق تعطى حرف جديد يليه .

٣- إن كل متوسطين يشتركان في حرف معين أو عدة حروف ليس بينها فرق معنوي (أي معاملتين تشتركان في حرف أو أكثر لا يوجد بينهما فرق معنوي) .

treat	C	B	D	A
mean	10	7	5	4
0.05	a	b	c	c
0.0[illegible]	a	b	b c	c

$\sqrt{11.62} = 3.٤١$

أولاً :اختبار L.S.D المسألة الهجن:

$$L.S.D_{0.05} = t_{0.05(21)} \cdot \bar{sd} = t_{0.05(21)} \cdot \sqrt{\frac{2S^2d}{ر}}$$

$$= 2.08 \cdot \sqrt{\frac{2(23.23)}{4}} = 7.09$$

$$L.S.D_{0.01} = 2.831 \cdot \sqrt{11.62} = 9.65$$

ثم ترتب المقارنات للحصول على اعتبار إن **A** هي معاملة الشاهد :

المقارنة COMPARISNS	الفرق بين المتوسطين Differences	L.S.D	
		0.05 7.09	0.01 9.65
B – A	13.75 – 16.5 = -2.75	-	-
C – A	17 - 16.5 = 0.50	-	-
D – A	8.50 – 16.5 = -8.00	*	-
E – A	19.25 - 16.50 = 2.75	-	-
F – A	16.75 - 16.50 = 0.25	-	-
G – A	25.5 – 16.50 = 9	*	-
H - A	21.25 – 16.50 = 4.75	-	-

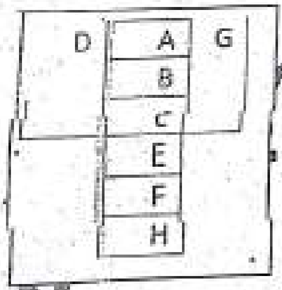
نلاحظ من خلال المقارنة السابقة :

١- وجود فروق معنوية بين الصنف D والصنف A:

٢- وجود فروق معنوية بين الصنف G والصنف A

٣- عدم جود فروق معنوية بين الصنف A وبقية الأصناف

والترتيب هو



ثانياً : اختبار دانكان لمسالة الهجن:

١- يحسب الخطأ المعياري :

$$S\bar{X} = \sqrt{\frac{S^2e}{r}} = \sqrt{\frac{23.23}{4}} = 2.41$$

٢- ترتيب المتوسطات تصاعدياً :(أو تنازليا) :

المعاملة	D	B	A	F	C	E	H	G
المتوسط	8.50	13.75	16.50	16.75	17.00	19.25	21.25	25.

٣- تستخرج قيم SSR من الجدول أمام درجات حرية الخطأ التجريبي وهي:**21**

P	2	3	4	5	6	7	8
SSR(0.05)	2.94	3.09	3.17	3.24	3.30	3.33	3.36
SSR(0.01)	4.00	4.19	4.30	4.38	4.44	4.50	4.56

٤- تحسب قيم (LSR) من المعادلة **LSR=SX̄ . SSR**

P	2	3	4	5	6	7	8
LSR 0.05	7.08	7.45	7.64	7.81	7.95	8.03	8.10
LSR 0.01	9.64	10.10	10.36	10.56	10.70	10.85	10.99

٥- يعمل جدول المقارنات على أن تبدأ باكبر المتوسطات فالأقل منه وهكذا :

G	H	E	C	F	A	B	D
25.5	21.25	19.25	17	16.75	16.5	13.75	8.5

إعداد

فخري العمر

P	المقارنة		LSR	
			0.01	0.05
8	G-D=17.00	**	8.10	10.99
7	G-B=11.75	**	8.03	10.85
6	G-A= 9	*	7.95	10.70
5	G-F=8.75	*	7.81	10.56
4	G-C=8.50	*	7.64	10.36
3	G-E=6.25	-	7.45	10.10
2	G-H=4.25	-	7.08	9.64
7	H-D=12.75	**	8.03	10.85
6	H-B=7.50	-	7.95	10.70
5	H-A=4.75	-	7.81	10.56
4	H-F=4.50	-	7.64	10.36
3	H-C=4.25	-	7.45	10.10
2	H-E=2.00	-	7.08	9.64
6	E-D=10.75	**	7.95	10.70
5	E-B=5.50	-	7.81	10.56
4	E-A=2.75	-	7.64	10.36
3	E-F=2.50	-	7.45	10.10
2	E-C=2.25	-	7.08	9.64

P	المقارنة		LSR	
			0.01	0.05
5	C-D=8.50	*	7.81	10.56
4	C-B=3.25	-	7.64	10.36
3	C-A=0.50	-	7.45	10.10
2	C-F=0.25	-	7.08	9.64
4	F-D=8.25	*	7.64	10.36
3	F-B=3.00		7.45	10.10
2	F-A=0.25		7.08	9.64
3	A-D=8.00	*	7.45	10.10
2	A-B=2.75	-	7.08	9.64
2	B-D=5.25	-	7.08	9.64

* تدل على إن الفرق بين المتوسطين معنوي عند مستوى 0.05

** تدل على إن الفرق بين المتوسطين معنوي عند مستوى 0.01

إعداد حمد العلي [illegible] العمر

٦- يلخص جدول المقارنات لنشره في الدوريات أو المجلات العلمية كما يلي :

المعاملة	G	H	E	C	F	A	B	D
المتوسط	25.00	21.25	19.25	17.00	16.75	16.50	١٣.٧٥	8.5
مستوى المعنوية 0..05	∂	∂b	∂bc	bcd	bcde	bcdef	bcdefg	g
مستوى المعنوية 0..01	∂	∂b	∂bc	∂bcd	∂bcde	∂bcdef	bcdefg	defg

ملاحظة : أي معاملتين تشتركان في حرف واحد أو أكثر ليس بينهما فرق معنوي عند مستوي المعنوية المطلوب ، ولتبسيط الجدول واختصار الحروف يمكن عمل الجدول بالطريقة التالية :

المعاملة	G	H	E	C	F	A	B	D
المتوسط	25.00	21.25	18.25	17.00	16.75	16.50	13.75	8.5
مستوي المعنوية 0.05	∂	ab	∂-c	b-d	b-e	b-f	b-g	g
مستوي المعنوية 0.01	[illegible]	∂b	∂-c	a-d	∂-e	∂-f	b-g	d-g