جامعة الموصل

كلية التربية البدنية وعلوم الرياضة

ملزمة الإحصاء للمرحلة الثانية

أ.م.عائدة يونس ال مراد   م.م. زياد يحيي علي

٢٠٢٢/٢٠٢١

University of Mosul, College of physical education and sport sciences, Mosul, Iraq.
Email: zead77@uomosul.edu.iq

الفصل الأول

مقدمة ووصف البيانات Introduction and data Description

1-1 مقدمة للتعريف بعلم الإحصاء:

استخدم الإحصاء قديما من قبل المصريين وطبقها الفراعنة في بناء الأهرامات حيث قاموا بالتعداد لسكان مصر وثروتها واستخدموا النتائج في تنظيم مشروع البناء، وكذلك في عصر الدولة الإسلامية استخدم الخليفة المأمون فكرة الحصر والعد لمعرفة عدد السكان ومقدار الزكاة وكان استخدام الإحصاء في البداية مقصوراً على الإعمال الخاصة بشؤون الدولة كما يدل على ذلك الأصل اللغوي في اسم هذا العلم وهو "**Statistics**" حيث كلمة "**State**" تعني الدولة المنشقة من كلمة "**Statistics**" أي تعني مجموعة من الحقائق الخاصة بشؤون الدول.

اما معنى إحصاء **لأي فرد فيقتصر على الجداول العددية التي تصف ظاهرة معينة او على الرسوم البيانية او الأشكال التصويرية التي تعرض التغيرات في ظاهرة خلال فترة زمنية معينة.** ويمكن ملاحظة ذلك من خلال حياتنا اليومية مثلاً التطلع في الصحف اليومية حيث يمكن ان تشاهد بعض الجداول التي تبين معدل كميات نزول المطر او تمثيل بيانات عن اسواق النقد والتقلبات في العملة والأسهم ... الخ من المعلومات و الرسومات الإحصائية.

ومن هذا يمكن أعطاء وصف صريح للإحصاء :

فهو علم كغيره من العلوم الأخرى له نظرياته وقوانينه وأساليبه ، وان علم الإحصاء هو العلم الذي يهتم بوصف طرق متعددة لجمع البيانات والمشاهدات ومن ثم يتم تنظيمها وعرضها باستخدام الأساليب العلمية لتحليلها واستخلاص النتائج منها، ولعلم الإحصاء علاقة وطيدة بمختلف العلوم الأخرى منها الرياضيات، العلوم الإنسانية، علم الاجتماع والدراسات السكانية، وكذلك العلوم الطبية والهندسية وعلم الأرض والحياة والوراثة وغيرها من العلوم.

وكان الإحصاء له دور كبير في حركة البحث العلمي في مجال الرياضة والتربية البدنية وعلى شتى فروعها كالصحة واللياقة البدنية والإدارة الرياضية والترويح والتدريب الرياضي وفسيولوجيا الحركة والميكانيكي الحيوية والمناهج وطرق التدريس والاجتماع الرياضي وعلم النفس الرياضي، والاختبارات والمقاييس والتقويم الرياضي وغيرها.

ويعتمد الباحثين بشكل واسع على دراسة العينات ليحصل على البيانات الإحصائية التي تتعلق بنشاطات الإنسان والأحوال المتعلقة به، وان كثير من الطرق الإحصائية قد تم استعمالها وتستعمل في الوقت الحاضر لجمع وتحليل وعرض البيانات لغرض التخطيط في كافة المجالات واتخاذ القرارات لذا يلعب التحليل الإحصائي دوراً بارزاً في كثير من حقول النشاط الإنساني والذي يعتبر مفيداً جداً في تبادل المعلومات والوصول الى الاستنتاجات والاستدلالات في البيانات ومن ثم في الإرشاد الى التخطيط المنطقي واتخاذ القرار.

ويمكن تقسيم الإحصاء الى قسمين رئيسيين :

1 – الإحصاء الوصفي Descriptive Statistics : والذي يمثل الطرق الرقمية او الحسابية لجمع المعلومات والبيانات لتلخيصها واختصارها ومن ثم عرض المعلومات عن طريق الجداول والرسوم البيانية وغيرها.

2 – الإحصاء الاستدلالي Inferential Statistics: وهو ذلك الجزء من الإحصاء الذي يعنى بتحليل البيانات المتوفرة في العينة Sample كأساس تحليل البيانات الموجودة في المجتمع Population للتوصل الى

أســاليب التقدير والاختبار واتخاذ القرارات والتنبؤ او الاســتقراء، ويلعب هذا الجزء دوراً مهماً في تخطيط التجارب التي تجمع منها البيانات ومن ثم تصميمها.

لذا فان الإحصاء يهتم بطرق جمع البيانات وتمثيلها وعرضها " الإحصاء الوصفي" ومن ثم تحليها وتفسيرها التوصل إلى الاستنتاجات " الإحصاء الاستدلالي" .

1-2 مراحل البحث العلمي :

يعتمد البحث العلمي على الطرق والأدوات الإحصائية المختلفة والمستخدمة في أي من العلوم او أي مجال حسب طبيعة ونوع البحث ، ومن اهم مراحل البحث العلمي هي :

1- تحديد المشكلة :

يتم تحديد نوع المشكلة وفي أي مجال او في أي علم من العلوم التي ذكرت سابقاً التي تستحق البحث والتقصي فدور الباحث هنا كيف يتم اختيار المشكلة المناسبة لدراسة الظواهر الغريبة او المألوفة ، وذلك بالتعبير عنها بعلاقة ما بين المتغيرات Variables او المشاهدات Observations بحيث تمكن الباحث من أجراء التحليل الإحصائي او ألاستنتاجي.

2- استخدام الأساليب الإحصائية :

بعد تحديد المشكلة للدراسة والإلمام بجميع جوانبها يتم بعد تفسير الظواهر تفسيراً علمياً من جميع الجوانب والطرق التي سوف تستخدم لحل المشكلة على قرار هذا ليتمكن الباحث من اختيار البيانات الملائمة لها وفي أي جانب التوجه والتي يتم فيها مرحلة جمع البيانات.

3- مرحلة جمع المعلومات او البيانات :

تعتمد على بعض الأساليب الإحصائية في جمع المعلومات " البيانات" والتي سوف نتطرق لها لاحقاً بشكل مفصل.وهناك مصادر هما:

أ-مصادر تاريخية ب-مصادر مدانية

4- تحليل البيانات :

استخدام الطرق والأساليب الإحصائية المختلفة لتحليل البيانات المتوفرة في العينة.

5- استخلاص النتائج ووضع التوصيات :

الاستناد على التحليل الإحصائي لبيانات العينة لوصف النتائج حول المجتمع ومن ثم اقتراح حلول للمشاكل ووضع التوصيات المختلفة، ومن كل ما ورد يتضح ان علم الإحصاء يعتمد على المفاهيم والأسس العلمية التالية

- **المجتمع :** "يعرف المجتمع بأنه عبارة عن جميع المفردات التي يمكن ان يأخذها المتغير". ويرمز له بالرمز (N).

ويتم استخدام المجتمع بأسلوب الحصر الشامل ، والحصر الشامل يطلق على الأسلوب الذي يجمع فيه الباحث بيانات بحثه من جميع المفردات التي تشكل مجتمعه ومن كافة الحالات التي تنطبق عليها خصائص معينة دون ترك أي مفردة او حالة ، وتسمى هذه الطريقة أيضاً بالتعداد .

ففي المجال الرياضي يعد طلبة كلية التربية الرياضية في جامعات بغداد والموصل والبصرة وديالى وبابل والقادسية مجتمع البحث لطلبة كليات التربية الرياضية في جمهورية العراق ، ويعد طلبة كلية التربية الرياضية في جامعة الموصل مجتمع البحث لطلبة كلية التربية الرياضية بجامعة الموصل ويعد طلبة الصفوف الاولى في كلية التربية الرياضية بجامعة

الموصل مجتمع البحث لطلبة السنة الأولى في كلية التربية الرياضية في جامعة الموصل لان المنطوق هنا حدد المجتمع أي جميع مفردات المجتمع المخصص سواء أكان كليات ، أم كلية ، أم سنة دراسية .

* **مميزات استخدام المجتمع (الحصر الشامل)**

1– دقة النتائج المتحصل عليه والموثوق في كفاءتها نظراً لجمع البيانات من كل فرد شمله البحث من دون ترك مفردة أو حالة .

2– تجنب أخطاء التعميم التي تنتج من استخدام بيانات مأخوذة من عينة محددة من المجتمع وتطبيق نتائجها على المجتمع كله .

3– تتفادى هذه الطريقة الأخطاء الشائعة والناجمة في غيرها من الطرائق (طريقة العينة) خاصة خطأ التحيز وخطأ الصدفة .

* **عيوب استخدام المجتمع (الحصر الشامل)**

1– باهظ التكاليف ويحتاج إلى إمكانيات طائلة .

2– يستغرق وقتا طويلاً وتبذل فيه جهود كبيرة في جمع البيانات وتصنيفها .

3– يحتاج الى جهاز أداري وفني ضخم ومدرب للقيام به .

- **العينة**: " تعرف العينة بأنها ذلك الجزء من المجتمع الذي يجري اختيارها على وفق قواعد وطرائق علمية بحيث تمثل المجتمع تمثيلاً صحيحاً " . ويرمز لها بالرمز (n) بحيث ان $N>n$

ويلجأ الباحث الى استخدام العينات في البحث ليس للاستفادة من المعلومات المتعلقة بالعينة ذاتها لصالحها الخاص ، ولكنه يجمع المعلومات للوصول الى حقيقتين رئيستين هما :

- لاستعمالها لاختبار فرضيات عن المجتمع الأصلي الذي سحبت منه العينة .
- أو لاستعمالها من اجل استخلاص استنتاجات عن طبيعة ذلك المجتمع الأصلي.

وتسمى هاتان العمليتان بالاستدلال الإحصائي .

ويلجأ الباحث في تعميم نتائج العينة على المجتمع الأصل.ولطريقة العينات مزايا وعيوب نوجزها في الآتي :

* **مميزات استخدام العينة (أو لماذا نستخدم العينات في البحوث)**

- تكتفي بعدد محدود من المفردات وليس جميعها وفي ذلك توفير للنفقات واختصار للجهد.
- تتيح فرصة سريعة في الوصول الى نتائج الدراسة لدى مقارنتها بالنتائج المتحصل عليها في طريقة الحصر الشامل .
- لا تحتاج إلى عدد كبير من الباحثين أو جامعي البيانات أو من الإحصائيين لاستكمال خطواتها وإجراءاتها .
- تتيح للباحث إصدار حكم أكثر تعمقاً .
- تستخدم العينات لأنها تكون اقل عرضة للأخطاء من أسلوب الحصر الشامل .

- يعد استخدام العينات احد وسائل أثراء البحوث العلمية الحقيقية لانه لا يمكن تعميم نتائجها بالنسبة للمجتمع الذي تمثله هذه العينات .
- تستخدم العينات في الحالات التي يكون فيها اجراء الحصر الشامل مستحيلا من الناحية العلمية وخصوصا في المجتمعات اللانهائية ، أي المجتمعات التي لا يمكن حصرها .
- ان تحليل النتائج التي نحصل عليها من الحصر الشامل تحتاج الى وقت طويل وقد تضيع الحكمة من التعداد اذا ما انتظرنا حتى يتم التحليل ، ولكن نتائج العينة يمكن الوصول إليها في وقت سريع يمكننا من الاستفادة منها.

*** عيوب استخدام العينة**

تتعرض نتائج البحث بطريقة العينة لأنواع مختلفة من الأخطاء تتمثل في أخطاء التعميم التي تنتج من استخدام بيانات مأخوذة من عينة محدودة من المجتمع ، **وتطبيق نتائجها على المجتمع كله واهم تلك الأخطاء** :

1- **خطأ التحيز** : وينتج من عدة أسباب أهمها أعطاء جميع الوحدات في المجتمع فرصا متساوية في الاختيار (عدم تحقيق مبدأ الاختيار العشوائي) ويحدث ذلك حينما يأخذ الباحث عينته المختارة من فئة معينة لها خصائص مميزة عن المجتمع الكلي . مثلاً اختيار طلبة كلية التربية الرياضية في جامعة الموصل للتعبير عن مستوى اللياقة البدنية لطلبة الجامعة ، اذ يمكن الخطأ هنا بالتحيز لاختيار عينة متخصصة في ممارسة الأنشطة الرياضية لا تعبر عن حقيقة المستوى العام للياقة البدنية لطلبة جامعة الموصل .

وقد يتحيز الباحث في اختيار أفراد العينة من أشخاص معينين من أصدقائه أو معارفه أو أقاربه أو المحتكين به واللذين يعرفهم ويثق بهم ، وهو أمر لا يحقق لجميع أفراد المجتمع الأصلي فرصاً متساوية في الاختيار ، فضلاً عن انه غالباً ما يكون لهم صفات مشتركة تميزهم عن غيرهم من أفراد مجتمع البحث ويحدث التحيز أحياناً عند ميول المتطوعين كأفراد للعينة سواء ممن لهم هدف من هذا التطوع أم لا ، وهم غالباً ما يتميزون بصفات خاصة .

ويحدث التحيز أيضاً نتيجة لخطأ فني في اختيار العينة وقد يحدث التحيز نتيجة عدم دقة الإطار وكفايته كإغفال مفردات من المجتمع مثل اختيار العينة في أطار غير شامل لكافة مفردات المجتمع الأصلي ممن لهم مختلف الصفات ، مثلاً استخدام إحصائيات قديمة عن الطلبة أو اللاعبين ، أو شمول الألعاب الفردية دون الألعاب الجماعية أو الاقتصار على عينات من الحضر دون الريف .

2- **خطأ الصدفة** : يزداد احتمال ورود هذا الخطأ كلما صغر حجم العينة ، في حين يقل هذا الاحتمال كلما زاد حجم العينة واقترب من حجم المجتمع الأصلي ، حتى ينعدم هذا الخطأ نهائياً إذا تم اختيار كل مفردات المجتمع الأساسي أي انه لا مجال لخطأ الصدفة في طريقة الحصر الشامل.

هذا وقد تلعب الصدفة دوراً في هذا الخطأ الذي ينشأ نتيجة للفروق بين خصائص افراد العينة المختارة وافراد المجتمع الأصلي ، فمثلاً إذا أدت الصدفة الى اختيار شخص رياضي بارز أو شخص ضعيف جداً في اللياقة البدنية فان كلا منهم كفيل بأن يزيد الانحراف المعياري لمستوى اللياقة البدنية .

– أنواع العينات

للعينات أنواع مختلفة وردت تحت أسماء عديدة حسب الوصف الذي يحدد أسلوب اختيار العينة ، إلا ان هناك تقسيم متفق عليه تقريباً وهو تقسيمها حسب أسلوب اختيارها نذكر منها :

أولاً – العينات غير الاحتمالية :

وهي العينات التي تؤخذ بغير الأسس السليمة ، ومن عيوبها أنها لا تمثل مجتمع البحث تمثيلاً دقيقاً ، ومن ثم فان نتائجها لا تصلح للتعميم على المجتمع كله بالثقة المناسبة ومن أمثلة هذا النوع ، العينات العارضة والحصصية والعمدية .

ثانياً – العينات الاحتمالية :

وقد سميت بهذا الاسم لأننا في اختيار هذا النموذج من العينات نعرف احتمال اختيار أي فرد في العينة ولذلك نستطيع تطبيق نظرية الاحتمالات ، وهي ضرورية عندما يكون الغرض من البحث هو الحصول على نتائج دقيقة يمكن تعميمها على المجتمع ، إذ يتسنى لنا هنا قياس الأخطاء الناتجة عن عملية المعاينة والتحكم في هذه الأخطاء وهذا النموذج من العينات له أسس إحصائية ومن أمثلة هذا النوع (العينات العشوائية البسيطة والطبقية والمنتظمة والعنقودية وعينة المجموعات ، والمتعددة المراحل) .

1-3 القياسات Measurements

ووحدات القياس للبيانات Measurements Scale

التحليلات والحسابات الإحصائية تحتاج الى تحديد نوع المقياس بالنسبة للبيانات حسب التحليل الإحصائي.
وهذه البيانات او قيم المتغيرات هي عبارة عن قياسات Measurements تقيس تلك الظواهر او تمثل قيم تلك المتغيرات، ومن وحدات القياس المستخدمة :

1– وحدات القياس / الفئات : وهي تخص تسمية المفردات او تصنيفها حسب عدد من التصنيفات المختلفة مثل تصنيف حسب الجنس (ذكور – إناث) او التصنيف حسب الحالة الاجتماعية (متزوج ، غير متزوج) او تصنيف حسب الحالة الصحية (مريض او غير ذلك).

2– وحدات القياس ألمرتبه : عندما لا تصنف البيانات وفقاً لتطبيقات مختلفة بل هنالك ترتيب في قياسها حسب مؤشرات معينة عندئذ نقول ان القياس قد تم تحت وحدات القياس المرتبة فمثلاً حالة المريض يمكن ان ينصف (مريض قليلاً ، مريض كثيراً ، متعافياً) او في ذكاء الطفل يصنف (يفوق المتوسط او اقل المتوسط) وهكذا.

– وحدات القياس بالفترات الزمنية : تحت هذا النوع من وحدات القياس نستطيع إضافة الترتيب القياسات لتحديد المسافات بين نوعين القياسات وان القياس هنا يعتبر قياس كمي.

4- وحدات القياس للنسب : وحدات القياس هنا تتميز كحقيقة تساوي النسب ولذلك تساوي الفترات او تشابه وحدات القياس بالفترات السابقة ذكرها من حيث نبدأ النقطة الصفرية فمثلاً ان النقطة تعني عدم وجود قيمة معينة للمفردة وفي حالة الأوزان فان الفرق بين 100Kg و50Kg تمثل الفرق بين حالتين وحدات القياس بالفترات والنسب.

4-1 الخطوات الأساسية للطرائق الإحصائية في البحث العلمي

- جمع البيانات .
- تصنيف البيانات وتبويبها .
- تمثيل البيانات .
- تحليل البيانات .
- الحكم على البيانات المقارنة .

1- جمع البانات

يستخدم الباحث احد أدوات البحث او وسائل جمع البيانات مثل المقابلة والاستبيان والاختبار ، والملاحظة والقياس للحصول على البيانات المطلوبة ، ويمكن الحصول على البيانات بطريقة مباشرة او غير مباشرة .

أ . **الطريقة غير المباشرة** : مثلاً الحصول على مستويات الطلبة المتقدمين للقبول في كلية التربية الرياضية من سجلات الكلية او لجنة الاختبارات دون أجراء الاختبارات او الاتصال بعينة او مجتمع البحث .

ب . **الطريقة المباشرة (الميدانية)** : "ويقوم الباحث بجمع البيانات عن طريق الاتصال بمفردات المجتمع موضوع البحث مباشرة وذلك عن طريق الاستبيان او المقابلة او الملاحظة او الاختبار ويقوم الباحث بجمع بياناته في استمارة إحصائية تصمم خصيصاً لهذا الغرض".

ومن طرق جمع البيانات :

– **المقابلة الشخصية** : يقوم الباحث بتوجيه الأسئلة للأشخاص او المفردات وتدوين الإجابات في الأماكن المخصصة ولهذا الأسلوب مميزات حيث يمكن جمع اكبر قدر ممكن من الإجابة الصحيحة وكذلك تكون نسبة الاستجابة للمعلومات مرتفعة.

– **المراسلة** : يتم باستخدام الاتصال الهاتفي او البريدي بحيث يذكر الباحث بكتاب مرفق أغراض البحث ومدى أهمية التعاون لإنجاح البحث، اما الاتصال الهاتفي يستخدم بدلا من المواجهة الشخصية لتقليل التكاليف وتوفير الوقت.

– **التسجيل المباشر** : وهنا يقوم الباحث بالانتقال الى موقع العمل ومشاهدة الأشياء او الأشخاص مثلاً البحوث التي تجري على السير او المرور في منطقة معينة، ظاهرة التهجين لبعض السلالات النباتية او الحيوانية ، وعادة ما يتم جمع البيانات باستخدام الاستمارة الإحصائية.

ويتم جمع البيانات من المجتمع بطريقة الحصر الشامل او من عينة تمثل المجتمع .

2- تصنيف البيانات وتبويبها (Classification and Tabulation)

" قبل البدء بتصنيف البيانات وتبويبها يجب ان يراجع الباحث الاستمارات او الأرقام التي حصل عليها ، والغرض من ذلك هو اكتشاف الاستمارات التي بها بيانات متناقضة او نقص في الإجابة او وجود ارقام او قيم شاذة تحتاج الى مراجعة وتدقيق ، وهنا يجب أعادة الاختبار او الاستبيان ، او إلغاؤه في حالة عدم أمكانية التصحيح " .

أ . **التصنيف** : وفيه فرزالباحث البيانات الى رياضي وغير رياضي ، مدخن وغير مدخن، وذكور وإناث ، ولاعبي كرة القدم ولاعبي كرة السلة وخريج ابتدائية او متوسطة او ثانوية، او التصنيف حسب المهنة او الحالة الزوجية ، او الديانة ... الخ .

ب . **التبويب** : وفيه تفريغ البيانات في جداول تسهل قراءتها وتلخص معالمها مثل الجداول التكرارية ، والجداول التكرارية المزدوجة والمتعددة او الإشكال البيانية ومنها يتم التبويب على أساس البيانات المطلوبة وقد يكون التبويب كمياً او نوعياً، ويمكن ملاحظة تفاصيل عرض البيانات في فصل لاحق، وهناك أربعة أشكال رئيسية للتبويب هي

- **تبويب زمني** : عبارة عن تجميع البيانات في جداول على أساس ان كل مجموع منها تعود لوحدة زمنية معينة كاليوم ، الأسبوع ، الأشهر.
- **تبويب جغرافي** : عبارة عن تجميع البيانات في جداول على أساس كل مجموعة منها خاصة بوحدة جغرافية معينة.
- **تبويب كمي** : عبارة عن تجميع البيانات في جداول على أساس كل مجموعة منها خاصة بوحدة كمية معينة كوحدات الوزن ، الطول ، المسافة ، الحجم.
- **التبويب على أساس صفة معينة** عبارة عن تجميع البيانات في جداول على أساس ان كل مجموعة تشترك بصفة معينة، عنوان وظيفي ، الحالة الاجتماعية ، الاختصاص.

3- تمثيل البيانات او العرض البياني (Graphic presentation)

بعد تصنيف البيانات وتبويبها يتم عرضها بيانياً اما بإشكال بيانية او بجداول تكرارية مبوبة ويمكن ملاحظة تفاصيل العرض البياني في الفصول اللاحقة.

4- تحليل البيانات

بعد تبويب وتصنيف البيانات يحتاج بعض منها الى معالجة إحصائية باستخدام مقاييس التشتت او النزعة المركزية او العلاقات وفي هذه الحالة يختار الباحث الوسائل الإحصائية اللازمة لتحليل البيانات والحصول على قيم يحتاجها للحكم على المتغيرات التي حلل بياناتها .

5- الحكم على البيانات

يحتاج الباحث لمعرفة وجود العلاقات او الارتباطات بين المتغيرات او معرفة الفروق إحصائيا باللجوء الى مقارنة النتائج التي حصل عليها في معالجاته الإحصائية مع قيم جدوليه ثابتة مثل قيم (t) . وقيم (F) وقيم (r) وغيرها . عندها يتأكد الباحث من صحة الفروض او عدم صحتها على ضوء هذه المقارنات

1-5 طبيعة البيانات

الصفة التي تتغير من شخص الى اخر او من مفردة الى أخرى تسمى ظاهرة Observation او المتغير Variable ويرمز له (X) ولكل مفردة او مشاهدة منها ترمز بالرمز (xi) $i= 1,2,\ldots n$ فمثلاً عند دراسة أطوال مجموعة من الطلاب في جامعة الموصل فأننا نرمز لصفة الطول بالرمز X وطول أي طالب او طالبة بالرمز (xi) وهذه تسمى بالمشاهدة او المفردة Observation ومن هنا يمكن تقسيم المتغيرات Variables الى نوعين :

1- متغيرات وصفية او نوعية Qualitative Variables :

وهي تلك الظواهر او الصفات التي لايمكن قياسها مباشرة بالأرقام العددية مثل صفة لون العين (ازرق ، اخضر ، اسود...) الحالة الاجتماعية (غني ، متوسط ، فقير ، معدوم) وكذلك صفة الجنس ، الى غير ذلك من الصفات.

2- المتغيرات الكمية (Quantitative Variables)

وهي تلك الظواهر او الصفات التي يمكن قياسها مباشرة بأرقام عددية مثل صفة الطول والوزن والعمر ... الخ وتقسم المتغيرات الكمية الى قسمين هما :

أ . متغيرات كمية مستمرة (او متصلة) (Continuous Variables)

فالمتغير المستمر هو المتغير الذي تأخذ المشاهدة او المفردة فيه اية قيمة رقمية في مدى معين ، فلو فرضنا بان أطوال طلبة جامعة ما تتراوح بين 180 , 150.5 سم فتقول بأن أطوال الطلبة تتراوح بين 150.5 ، 180 سم أي ان المتغير يمكن ان يأخذ قيمة بين 150.5 سم و 180 سم . فنقول بان : $150.5 \leq x \leq 180$

وأمثلة أخرى مثل المسافة والارتفاع والزمن والوزن كلها يمكن قياسها بأجزاء صغيرة جداً وتأخذ أية قيمة تقع في حدود معينة ، وبصورة عامة فان كل البيانات التي تقاس تعد بيانات لمتغير مستمر .

ب . متغيرات كمية غير مستمرة (متقطعة) (Discrete Variables)

المتغير المنفصل هو المتغير الذي تأخذ المشاهدة او المفردة فيه قيماً متباعدة او متقطعة غير مستمرة ، مثل عدد الطلبة وعدد الكرات او عدد القاعات الدراسية والملاعب ، والمسابح متغيرات كمية غير مستمرة وبصورة عامة فان كل البيانات التي نحصل عليها من العد تعد بيانات لمتغير متقطع (غير مستمر). فنقول بان $x = 2,3,4,5$.

1-6 الاستمارة الإحصائية / استمارة استبيان :

تتألف عادة من صفحة او عدة صفحات تحتوي على الأسئلة التي يريد الباحث إجابات عنها يترك فراغ عن كل سؤال لتسجيل الإجابة عليها ، وهنا يختلف تصميم الاستمارة من باحث الى اخر وحسب موضوع البحث ، ألا ان هناك شروطا عامة يجب مراعاتها عند تصميم الاستمارة منها :

1 – ان تكون الأسئلة مختصرة وواضحة ولا تحتمل اكثر من تفسير.

2– ان يكون عددها اقل ما يمكن.

3– تفضل الأسئلة ذات الإجابة بنعم او لا او ذات الإجابات المتعددة و اختيار إجابة منها او ذات الإجابة المختصرة.

4– ان تكون الإجابات قابلة للتصنيف والتبويب ، ومن أسس التصنيف :

- اساس نوعي : (ذكور وإناث).
- اساس جغرافي : المناطق او المحافظات او حسب (ريف وحضر).
- اساس زمني : تصنيف المهاجرين حسب سنة دخولهم.
- اساس مشترك : حسب معيارين مثلاً (ريف –ذكور)(ريف – إناث)(حضر– ذكور) (حضر – إناث).

ويجب الإشارة الى انه من الضروري تجربة الاستمارة بعد تصميمها وذلك بهدف اكتشاف النواقص ونقاط الضعف فيها وتعديلها اذا لزم الأمر، كما انه ومن المهم الإشارة وبوضوح في بداية الاستمارة الى سرية المعلومات والى نوع وأهمية الدراسة التي تقوم بها، مما يشجع الوحدة الإحصائية الدقة في الإجاب

1-7 عرض البيانات الإحصائية :

البيانات الأولية Row data الخاصة بالدراسة ، لايمكن تفسيرها بشكل ملائم والاستفادة منها وهي بهذه الصورة لذلك يلجا الباحث الى وضع تلك البيانات بشكل جداول مبسطة او التعبير عنها بالرسوم البيانية لكي يسهل دراستها وتحليلها ومن طرق عرض البيانات نذكر :

هناك خمس طرائق لعرض البيانات الإحصائية هي:

١– العرض الكتابي	Text presentation
٢– العرض شبه الجدولي	Semi – tabular presentation
٣– العرض الجدولي	Tabular presentation
٤– العرض الرتبي	Range order
٥– العرض البياني	Graphical presentation

- **العرض ألجدولي Tabular Presentation** : وهي عبارة عن عرض البيانات التي تم جمعها و وصفها في جداول منتظمة لذا فهناك نوعان من الجداول .

١- الجداول البسيطة Simple Table: وهي تلك الجداول التي تتوزع فيها البيانات حسب صفة واحدة ويتألف عادة من عمودين ، الأول يمثل الظاهرة Observation والثاني يمثل عدد المفردات التابعة لكل مشاهدة وما يسمى بالتكرارات Frequency ويرمز لها بالرمز fi، ومن هذا المنطق يمكن تسمية الجداول بجداول التوزيع التكراري Frequency Distribution Table .

اما الطريقة الأساسية لبناء التوزيع التكراري فهي عبارة عن تقسيم مدى قيم البيانات والتي تمثل المجموعة الكبيرة الى فئات تمثل مجموعات صغيرة مختلفة ثم يتم حصر عدد البيانات الواقعة ضمن كل فئة، ويمكن استخدام جداول التوزيع التكراري في حالة البيانات الوصفية وكذلك في حالة البيانات الكمية المتقطعة

مثال (1) : في حالة البيانات الوصفية

كون جدول تكراري للبيانات التالية والتي تمثل التوزيع الجغرافية لـ 35 طالباً من المقبولين في كلية التربية الرياضية لهذا العام .

موصل	**كركوك**	**بغداد**	**اربيل**	**اربيل**	**موصل**	**موصل**
موصل	**كركوك**	**كركوك**	**اربيل**	**موصل**	**بغداد**	**بغداد**
اربيل	**أربيل**	**موصل**	**كركوك**	**اربيل**	**موصل**	**موصل**
اربيل	**كركوك**	**اربيل**	**موصل**	**موصل**	**موصل**	**بغداد**
موصل	**كركوك**	**بغداد**	**كركوك**	**كركوك**	**كركوك**	**موصل**

الحل :

نكون اولاً جدولاً مكون من ثلاث أعمدة العمود الأول يمثل الظاهرة Observation والعمود الثاني لتفريغ البيانات اما العمود الثالث فيمثل التكرار لكل مشاهدة او صفة وما يسمى Frequency. وبذلك فان جدول التوزيع التكراري للموقع الجغرافي سيكون كالأتي

التكرارات	تفريغ البيانات	الموقع الجغرافي(محافظة)
١٣	IIII IIII IIII I	موصل
٩	I IIII IIII	كركوك
٥	I IIII	بغداد
٨	IIII IIII	اربيل
٣٥		المجموع

التوزيعات التكرارية Frequency Distribution : <u>في حالة البيانات الكمية</u>

عندما يكون عدد المفردات قليلاً او صغيراً يمكن عندئذ ملاحظة كل منها على انفراد اما عندما يكون عدد المفردات كبيراً فمن الأسهل والأفضل استخدام ما يسمى بالجدول التكراري Frequency . ويتم

ذلك بتجميع او توزيع البيانات كمجموعة كبيرة من القيم الى اكثر من مجموعة او مايسمى بالفئة Class او المجال Interval .

ولبناء جدول التوزيع التكراري علينا ملاحظة ما يلي :

1- تحدي المدى Range :

والذي يمثل الفرق بين اكبر واصغر قيمة في المجموعة وكما بالمعادلة الآتية :

المدى (R) = أعلى قيمة – أدنى قيمة +١

$$R = XU - XL + 1$$

2- اختيار عدد الفئات Number of Classes :

يمكن القول انه ليست هناك قاعدة عامة ولكن يمكن اختيار ذلك العدد من الفئات والذي يتناسب مع حجم البيانات والأهداف المتوخاة من التحليل ويعتمد ايضاً على خبرة الباحث ، ويمكن ان نختار عدد الفئات فرضياً على ان لاتقل عن 5 والا تزيد عن 15 وذلك تبعاً لطبيعة البيانات وعدد مفرداتها.

او من ممكن استخدام قانون او طريقة سترجس وهي كالاتي :

$$m = 1 + 3.3 \log (n)$$

3- تحديد ما يسمى بطول الفئة او مداها Class Width :

ويعتمد ذلك على خبرة الباحث وهيئة البيانات، وبشكل عام نستخدم عرض يسهل الحسابات عليها كأن نستخدم عرض فئة 5 او 15 وحدة وهكذا. ويمكن الاستناد على العلاقة التالية لتحديد ذلك العدد كالأتي:

$$Classeswidth(w) = \frac{Range}{Numberofclasses}$$ (مقربة الى اقرب عدد صحيح اكبر)

4- حدود الفئات Class Limits :

تبدأ حدود الفئات بأصغر قيمة او اصغر من ذلك بقليل والتي تمثل الحد الأدنى للفئة الأولى وتنتهي الفئة الأخيرة بالحد الأعلى والذي يمثل اكبر قيمة او اكبر من ذلك بقليل وبذلك فان كل فئة لها حدين هما الحد الأدنى Lower Cut Point ويرمز له (L) والحد الأعلى Upper Cut Point

وتعبر هذه الصيغة من انسب الصيغ المستخدمة لكتابة الفئات وتعني ان الحد الأدنى يدخل ضمن الفئة المعينة بينما لا يدخل الحد لأعلى فيها وإنما سيدخل في الفئة التي تليها، فانه يمكن تكوين حدود الفئات على النحو التالي :

أ- في حالة المتغيرات المتقطعة

الحد الأعلى للفئة	الحد الأدنى للفئة	تسلسل الفئة
XL+w-1	XL	1
XL+2w-1	XL+w	2
XL+3w-1	XL+2w	3
		.
		.
XL+nw-1	XL+(n-1)w	n

ب- في حالة المتغيرات المستمرة

الحد الأعلى للفئة	الحد الأدنى للفئة	تسلسل الفئة
XL+w	XL	1
XL+2w	XL+w	2
XL+3w	XL+2w	3
		.
		.
XL+nW	XL+(n-1)w	N

5- اما عن مراكز الفئات Class Midpoint or Class Marks

فأنها تمثل المتوســـط للحد الأعلى والحد الأدنى لكل فئة أي ان مركز الفئة والذي يرمز له بالرمز x_i يمكن حسابه باستخدام العلاقة :

$$X_i = \frac{Ui + Li}{2}$$

6- تفريغ البيانات وايجاد عدد التكرارات لكل فئة Class frequency :

تشكيل الجدول التكراري يعني توزيع المشاهدات او البيانات الموجودة على العدد المحدد المناسب من الفئات ومن ثم تحديد عدد المشاهدات او البيانات المفرغة في كل فئة.

يتم تسجيل القيم الواحدة بعد الأخرى في الفئة الخاصة بها على شكل أشارة او علامات او لا ثم ترجمتها الى أعداد (كما سيتم توضيح ذلك لاحقاً) وبعدها يتم جمع هذه التكرارات للتأكد من المجموع والذي يمثل حجم العينة ، العدد المعين من المشاهدات او البيانات والذي يخص فئة معينة يسمى بتكرار الفئة Class frequency ويرمز له (f_i) والذي يمثل تكرار الفئة i .

وبذلك فان الشـــكل العام للجدول التكراري يتكون من (3) أعمدة رئيســية وهي عمود للفئات وعمود للقيم الموزعة (التكرار بالإشارات او العلامات) وعمود يمثل مجموعة عدد القيم (التكرار) ، ويضاف عادة عمود اول يشير الى تسلسل الفئات.

سوف نتطرق الى التوزيع التكراري في حالة البيانات المتقطعة والبيانات المستمرة بأخذ مثال لتوضيح كل حالة.

اولاً : البيانات المتقطعة Discrete Data :

وهي تلك البيانات التي تأخذ قيمها أرقام عددية صحيحة مثل البيانات الخاصة بأعداد الأسرة او أعداد المرضى... الخ من البيانات ولتبويب هذه البيانات يجب تصنيفها في مجموعات تصاعدياً او تنازليا ثم وضع التكرار مقابلها.

والمثال التالي يوضح ذلك :

مثال (2) : البيانات التالية تمثل درجات (45) طالباً من طلبة جامعة الموصل/كلية التربية في مادة الإحصاء، كون جدول توزيع تكراري.

76	75	95	96	92	84	80	64	65	59	52	50	44
$\boxed{40}$	70	68	65	66	76	77	$\boxed{98}$	84	67	57	56	49
54	45	70	58	76	64	80	90	90	87	86	74	69
72	73	88	66	85	89							

الحل : لعمل الجدول التكراري نتبع الخطوات التالية :

1 . المدى $R = 98 - 40 + 1 = 59$

2 . عدد الفئات كما ذكرنا سابقاً ونفرض عدد الفئات تساوي 6.

او استخدام طريقة سترجس

$$m = 1 + 3.3\log(n) = 6$$

3 . إيجاد طول الفئة

$$w = \frac{R}{m}$$

وللحفاظ على توازن الفئات قربنا طول الفئة الى 10.

$$w = \frac{59}{6} \approx 10$$

4 . حدود الفئة : بما ان المطلوب في صياغة السؤال تكوين جدول توزيع تكراري لدرجات طلاب لذلك سوف يكون المتغير من النوع المتقطع بذلك سوف نستخدم الجدول الاول الخاص بهذا النوع من المتغيرات وكما يلي :

الحد الأعلى للفئة	الحد الأدنى للفئة	تسلسل الفئة
XL+w-1 40+11-1=(49)	XL (40)	1
XL+2w-1 40+2(10)-1= (59)	XL+w 40+10= (50)	2
XL+3w-1 40+3(10)-1 =(69)	XL+2w 40+2(10)= (60)	3
XL+4w-1 40+4(10)-1 =(79)	XL+3w 40+3(10)= (70)	4
XL+5w -1 40+5(10)-1 =(89)	XL+4w 40+4(10)= (80)	5
XL+6w-1 40+6(10)-1 =(99)	XL+5w 40+5(10)= (90)	6

5. أيجاد مركز الفئات x_i .

6. تفريغ البيانات في الجدول ثم نجد التكرارات المقابلة لكل فئة.

وأخيرا فان جدول التوزيع التكراري لعلامات الطلبة سيكون

جدول (2) التوزيع التكراري لدرجات طلبة جامعة الموصل/كلية التربية

التكرار (fi)	التكرار بالإشارات	مراكز الفئات (xi)	الحدود الفعلية للفئات	الفئات	تسلسل الفئة
4	IIII	44.5	40 - 49	40	1
7	III IIII	54.5	50 - 59	50	2
9	IIIII IIII	64.5	60 - 69	60	3
10	IIIII IIIII	74.5	70 – 79	70	4
9	IIIII II I	84.5	80 – 89	80	5
6	IIIII I	94.5	90 - 99	90- 100	6
45				المجموع	

واجب : اوجد جدول التوزيع التكراري لاختبار رمي الكرة الطبية لـ (99) طالب

القيم الآتية تمثل نتائج اختبار رمي الكرة الطبية

24	27	28	29	30	31	32	33	34
35	36	37	38	39	40	41	42	42
42	43	43	43	43	44	44	44	44
44	45	45	45	45	45	45	46	46
46	46	47	47	47	47	47	47	47
48	48	48	48	48	48	49	49	49
49	49	49	49	50	50	50	50	50
50	50	50	51	51	51	51	51	51
51	51	52	52	52	52	53	53	53
53	53	53	53	53	54	54	54	54
55	55	55	56	56	56	57	58	59

الحل : لعمل الجدول التكراري نتبع الخطوات التالية :

حتى نستطيع الحصول على الحل بصورة سريعة وصحيحة في البداية نرتب البيانات من اصغر قيمة الى اعلى قيمة (ترتيب تصاعدي)، وسوف نحصل على ما يأتي :

24	**35**	**42**	**44**	**46**	**48**	**49**	**50**	**51**	**53**	**55**
27	**36**	**43**	**45**	**46**	**48**	**49**	**50**	**51**	**53**	**55**
28	**37**	**43**	**45**	**47**	**48**	**49**	**50**	**52**	**53**	**55**
29	**38**	**43**	**45**	**47**	**48**	**49**	**51**	**52**	**53**	**56**
30	**39**	**43**	**45**	**47**	**48**	**50**	**51**	**52**	**53**	**56**
31	**40**	**44**	**45**	**47**	**48**	**50**	**51**	**52**	**54**	**56**
32	**41**	**44**	**45**	**47**	**49**	**50**	**51**	**53**	**54**	**57**
33	**42**	**44**	**46**	**47**	**49**	**50**	**51**	**53**	**54**	**58**
34	**42**	**44**	**46**	**47**	**49**	**50**	**51**	**53**	**54**	**59**

1 . المدى (R) = أعلى قيمة – أدنى قيمة

المدى $R = 59 - 24 + 1 = 36$

2 . **عدد الفئات كما ذكرنا سابقاً ونفرض عدد الفئات تساوي 8 كون حجم العينة كبير.**

اما في بعض المصادر فعملية ايجاد عدد الفئات تستخدم عدة طرق ومن هذه الطرق طريقة سترجس

$$w = \frac{R}{m}$$

$$w = \frac{36}{8} \approx 5$$

$$m = 1 + 3.3\log(n) = 1 + 3.3\log(99) = 1 + 3.3 * 1.995635 = 7.585596 \approx 8$$

3 . إيجاد طول الفئة

سوف نختار عدد الفئات يساوي (8) وطول الفئة يساوي (5)

ملاحظة هامة جداً : بامكاننا استخدام الطريقة المتبعة وهي ان نختار عدد الفئات ما بين 5 الى 15 او طريقة سترجس وذلك حسب السؤال وحجم العينة كما اشرنا .

4 . حدود الفئة : بما ان المطلوب في صياغة السؤال تكوين جدول توزيع تكراري لاختبار رمي الكرة الطبية لـــ(99) طالب لذلك سوف يكون المتغير من النوع المتقطع بذلك سوف نستخدم الجدول الاول الخاص بهذا النوع من المتغيرات وكما يلي :

تسلسل الفئة	الحد الأعلى للفئة	الحد الأدنى للفئة	الحدود الفعلية للفئات	مراكز الفئات (xi)	التكرار بالإشارات	التكرار (fi)
1	XL+w−1 24+5−1=(28)	XL (24)	24 — 28	26	III	3
2	XL+2w−1 24+2(5)−1= (33)	XL+w 24+5= (29)	29 — 33	31	IIIII	5
3	XL+3w−1 24+3(5)−1 =(38)	XL+2w 24+2(5)= (34)	34 — 38	36	IIIII	5
4	XL+4w−1 24+4(5)−1 =(43)	XL+3w 24+3(5)= (39)	39 — 43	41	IIIII IIIII	10
5	XL+5m−1 24+5(5)−1 =(48)	XL+4w 24+4(5)= (44)	44 — 48	46		28
6	Xs+6w−1 24+6(5)−1 =(53)	XL+5w 24+5(5)= (49)	49 — 53	51		35
7	XL+7w−1 24+7(5)−1 =(58)	XL+6w 24+6(5)= (54)	54 — 58	56		12
8	XL+8w−1 24+8(5)−1 =(63)	XL+7w 24+7(5)= (59)	59 — 63	61	I	1
المجموع						99

5. أيجاد مركز الفئات x_i .

6. تفريغ البيانات في الجدول ثم نجد التكرارات المقابلة لكل فئة

ثانياً : البيانات المستمرة Continuous Data : وهي تلك البيانات التي يمكن ان تأخذ ارقاماً صحيحة وكسرية والأمثلة كثيرة على هذا الجانب مثلاً أطوال مجموعة من الطلبة ، أوزانهم ، أعمارهم ، المنتجات بالطن، المبيعات ، كميات الإنتاج وغيرها من المتغيرات المتصلة.

ولبناء جدول توزيع تكراري يجب مراعاة الخطوات السابقة التي ذكرت والتي يمكن توضيحها من خلال المثال التالي :

مثال (3) : القيم التالية تمثل أطوال 80 طالب من طلاب كلية التربية الرياضية /جامعة الموصل (مقربة الى اقرب سنتمتر) والمطلوب أنشاء جدول توزيع تكراري لأطوال هؤلاء الطلاب.

180	187	198	181	174	148	179	180
178	182	193	191	170	190	180	184
173	174	181	156	165	192	170	171
186	183	193	165	151	185	168	172
168	186	143	174	173	183	190	135
175	167	172	190	171	176	192	193
181	188	191	197	172	161	180	191
177	171	159	180	195	199	170	174
163	189	167	160	182	183	163	160
175	179	188	166	170	188	176	163

الحل : **نتبع الخطوات الآتية** : حتى نستطيع الحصول على الحل بصورة سريعة وصحيحة ، في البداية نرتب البيانات من اصغر قيمة الى اعلى قيمة (ترتيب تصاعدي) ، وسوف نحصل على ما يأتي :

191	187	181	178	173	170	163	$\boxed{135}$
192	188	182	179	174	170	163	143
192	188	182	179	174	170	165	148
193	188	183	180	174	171	165	151
193	189	183	180	174	171	166	156
193	190	183	180	175	171	167	159
195	190	184	180	175	172	167	160
197	190	185	180	176	172	168	160
198	191	186	181	176	172	168	161
$\boxed{204}$	191	186	181	177	173	170	163

1. المدى (R) = أعلى قيمة – أدنى قيمة + ١

المدى R =204-135+1=70

2 . عدد الفئات كما ذكرنا سابقاً ونفرض عدد الفئات تساوي 7.

او نستخدم طريقة سترجس وهي كالأتي :

$$m = 1 + 3.3\log(n) = 1 + 3.3\log(80) = 1 + 3.3*1.90309 = 7.280197 \approx 7$$

3 . إيجاد طول الفئة

$$w = \frac{Range}{m}$$

$$w = \frac{70}{7} = 9.1428 \approx 10$$

وللحفاظ على توازن الفئات قربنا طول الفئة الى 10.

4. أيجاد مركز الفئات x_i .

6. تفريغ البيانات في الجدول ثم نجد التكرارات المقابلة لكل فئة.

وأخيرا فان الجدول (3) التوزيع التكراري لأطوال 80 الطلبة سيكون :

التكرار بالإشارات	التكرار (fi)	مراكز الفئات (xi)	الحدود الفعلية للفئات	الفئات	الحد الأدنى للفئة	الحد الأعلى للفئة	تسلسل الفئة
II	**2**	**140**	**135 – 145**	**135–**	**XL (135)**	**XL+w (145)**	**1**
II	**2**	**150**	**145 – 155**	**145–**	**XL+w (145)**	**XL+2w (155)**	**2**
IIII IIII	**8**	**160**	**155 – 165**	**155–**	**XL+2w (155)**	**XL+3w (165)**	**3**
	23	**170**	**165 – 175**	**165–**	**XL+3w (165)**	**XL+4w (175)**	**4**
	22	**180**	**175 – 185**	**175–**	**XL+4w (175)**	**XL+5w (185)**	**5**
	19	**190**	**185 – 195**	**185–**	**XL+5w (185)**	**XL+6w (195)**	**6**
IIII	**4**	**200**	**195 – 205**	**195–**	**XL+(n−1)w (195)**	**XL+nL (205)**	**7**
	المجموع 80						**n=7**

واجب : البيانات التالية تمثل أوزان 50 طالب في إحدى الكليات

49 54 19 38 44 24 46 43 57 29 47

56 4937 41 47 45 53 29 34 49 43 28

45 42 52 51 32 31 29 34 37 18 21

39 28 45 42 2228 37 32 27 26 41 39

43 35 **23** 47

المطلوب : كون جدول توزيع تكراري

2- **الجداول الثنائية** : وهي تلك الجداول التي تتوزع فيها البيانات حسب صفتين او ظاهرتين او أكثر في نفس الوقت، وبذلك فان الجدول الثنائي يتكون من :

الصـــفوف والتي تمثل فئات او مجاميع إحدى الظاهرتين والأعمدة التي تمثل فئات او مجاميع الصـــفة الأخرى الى المربعات الناتجة من الصفوف والأعمدة فتحتوي على عدد المفردات او التكرارات المشـتركة بين الظاهرتين، والمثال التالي يبين وصف الجداول الثنائية.

مثال (4) : البيانات التالية تمثل الطلبة المقبولين للعام الدراسي 1986/1985 لـ (50) طالباً حسب الجنس

جدول رقم (4) تمثل الطلبة المقبولين للعام الدراسي 1986/1985 لـ (50) طالباً حسب الجنس

التحصيل الإعدادي / الجنس	إعدادية	صناعة	تجارة	اولمبية	المجموع
ذكور	12	7	5	4	28
إناث	8	1	7	6	22
المجموع	20	8	12	10	50

مثال (5) : جدول التوزيع التكراري الثنائي التالي يمثل عدد من طلبة جامعة الموصل كلية التربية الرياضية مصنفين حسب صفتي الطول والوزن

جدول رقم (5) الطلاب حسب الوزن والطول

الطول \ الوزن	50	60	70-80	المجموع
120-	12	4	-	16
140-	3	25	6	34
160-180	5	6	14	25
المجموع	20	35	20	75

3- التوزيع التكراري النسبي Relative Frequency distribution

وهو جدول يبين الأهمية النسبية لكل فئة او مراكز الفئة مع التكرار النسبي يسمى جدول تكراري نسبي،ويمثل

قسمة تكرار كل فئة الى مجموع التكرارات : $\mathrm{Re}\,l.\,fi = \dfrac{fi}{\sum fi = n}$

وبذلك فان الجدول الذي يعرض الفئة او مراكز الفئة مع التكرار النسبي يسمى جدول تكراري نسبي وعليه يمكن حساب التكرار النسبي للمثال رقم (1) ، اذ ضربنا التكرار النسبي

* 100% يسمى بالتكرار النسبي المئوي **ويجب ملاحظة ان مجموع التكرارات النسبية دائماً يساوي واحد**

مثال (6) : احسب التكرار النسبي والنسبي المئوي للأجر الأسبوعي لـ(50) عاملاً

جدول رقم (6) للأجر الأسبوعي لـ(50) عاملاً

الفئات	التكرار (fi)	Rel.fi	Rel. f_I
40-	5	5/50	5/50%
50-	9	9/50	9/50%
60-	9	9/50	9/50%
70-	10	10/50	10/50%
80-	11	11/50	11/50%
90- 100	6	6/50	6/50%
المجموع	50	1	

4- التوزيع التكراري المتجمع Cumulative frequency distribution

سبق وتحدثنا عن الجداول التوزيع التكراري للفئات الذي يبين توزيع قيم الظاهرة، ولكن في بعض الأحيان قد نحتاج الى معرفة عدد القيم او المفردات (المشاهدات) التي تقل او تزيد عن قيمة معينة، لذا سوف نتطرق الى جانب التوزيعات التكرارية المتجمعة للفئات وهناك نوعان من هذه التوزيعات المتجمعة وهما:

أ- التوزيع التكراري المتجمع الصاعد ^Increasing Cumulative frequency CF

ويرمز له عادة اما ucf او ^CF وجدول التوزيع التكراري المتجمع الصاعد يتكون عادة من عمودين، العمود الأول يمثل الحدود والعمود الثاني التكرار المتجمع الصاعد الذي يكون حسابه كما يلي :

1- التكرار المتجمع الصاعد للفئة الأولى يمثل نفس تكرار الفئة الأولى لان تكرار الفئة السابقة للفئة الأولى = صفر ، أي ان :
$$F_1 \uparrow = f_1$$

2- التكرار المتجمع الصاعد للفئة الثانية يساوي التكرار المتجمع للفئة الأولى + تكرار الفئة الثانية أي :
$$F_2 \uparrow = F_1 \uparrow + f_1$$

3- هكذا نستمر بإضافة تكرار الفئة التالية للتكرار المتجمع الصاعد الى ان نصل الى اخر فئة حيث يمثل مجموع التكرارات الكلية أي ان :
$$F_n \uparrow = \sum_{i=1}^{n} f_i$$

كذلك يمكن استخراج التكرار المتجمع الصاعد النسبي لكل فئة وذلك بقسمة التكرار التجميعي الصاعد لكل فئة على مجموع التكرارات بحيث يكون التكرار النسبي الصاعد للفئة الأخيرة يساوي واحد.

مثال (7) : في الجدول التالي يبين التكرار المتجمع الصاعد والتكرار المتجمع النسبي باستخدام نفس المثال السابق.

جدول رقم (7) التكرار المتجمع الصاعد للأجر الأسبوعي والتكرار المتجمع الصاعد النسبي لـ(50) عاملاً

الفئات	التكرار (fi)	حدود العليا	CF_i^	Rel - CF_i^
40-	5	اقل50	5	5/50
50-	9	اقل60	14	14/50
60-	9	اقل70	23	23/50
70-	10	اقل80	33	33/50
80-	11	اقل90	44	44/50
90- 100	6	اقل100	50	50/50
المجموع	50			

مثال (8) : في الجدول الأتي والذي يمثل الحدود الحقيقية للفئات في اختبار رمي الكرة الطبية

الجدول المرقم (8) التكرار المتجمع الصاعد والتكرار المتجمع الصاعد النسبي لاختبار رمي الكرة الطبية

الفئات	التكرار (fi)	حدود العليا	CF_i^	Rel - CF_i^
24-30	0	اقل30	0	0/100
31- 37	5	اقل37	5	5/100
38- 44	7	اقل44	12	12/100
45- 51	16	اقل51	28	28/100
52- 58	47	اقل58	75	75/100
59- 65	25	اقل65	100	100/100
المجموع	100			

ملاحظة : في المثال رقم (7) البيانات مستمرة والمثال رقم (8) البيانات متقطعة ولا يوجد اختلاف في ايجاد التكرار المتجمع الصاعد لان الحل سوف يعتمد على التكرار وليس على حدود الفئة فقط في حالة الرسم الهندسي نعتمد على حدود الفئة .وكذلك الحال في ايجاد التكرار المتجمع النازل.

ب- التوزيع التكراري المتجمع النازل $CF_i \vee$ Decreasing Cumulative frequency :

ويرمز له ايضاً LCF او $CF_i \vee$ وهو التوزيع الذي يعطي عدد المفردات التي تزيد قيمها عن الحد الأدنى لفئة معينة ، ويتكون هذا التوزيع من عمودين، عمود يمثل الفئات والثاني يمثل التكرار المتجمع النازل ، والذي يمكن حسابه كما يلي :

1- التكرار المتجمع النازل للفئة الأولى يتمل بمجموع التكرارات الكلية أي ان : $F_1 \downarrow = \sum_{i=1}^{n} f_i$

2- التكرار المتجمع النازل للفئة الثانية يســـاوي التكرار المتجمع النازل للفئة الأولى مطروحاً منه التكرار المناظر له أي تكرار الفئة الأولى أي ان : $F_2 \downarrow = \sum_{i=1}^{n} f_i - f_1$

3- التكرار المتجمع النازل للفئة الثالثة سوف يكون : $F_3 \downarrow = F_3 \downarrow - f_2 = \sum_{i=1}^{n} f_i - f_1 - f_2$

4- وهكذا نستمر بالتنازل الى اخر فئة للحصـول على التكرار المتجمع النازل له F_n يسـاوي تكرار الفئة الأخيرة f_n ، وكذلك الحال بالنسبة الى التكرار المتجمع النازل النسبي يشبه التكرار النسبي الصاعد في عملية استخراجه .

مثال (9) : يبين التكرار المتجمع النازل والتكرار التجميعي النسبي وكذلك التجميعي الصاعد .

جدول رقم (9) التكرار المتجمع الصاعد والنازل للأجر الأسبوعي والتكرار المتجمع الصاعد والنازل النسبي لـ(50) عاملاً

Rel- $CF_i \vee$	Rel - CF_i^	$CF_i \vee$	حدود دنيا	CF_i^	حدود العليا	التكرار (fi)	الفئات
50/50	5/50	50	40 فاكثر	5	50اقل	5	40-
45/50	14/50	45	50 فاكثر	14	60اقل	9	50-
36/50	23/50	36	60 فاكثر	23	70اقل	9	60-
27/50	33/50	27	70 فاكثر	33	80اقل	10	70-
17/50	44/50	17	80 فاكثر	44	90اقل	11	80-
6/50	50/50	6	90 فاكثر	50	100اقل	6	90- 100
						50	المجموع

ج – جداول التوزيع التكراري المتجمع الصاعد والنازل :

يمكن دمج التكرار المتجمع النازل والصاعد بنفس البيانات والمطلوب ونفس التمثيل البياني وكما موضح في المثال أعلاه .

واجب : البيانات التالية تمثل الأجر الأسبوعي لـ (40) عاملاً بالدينار العراقي

المطلوب :

1- اوجد التكرار المتجمع الصاعد.

2- اوجد التكرار المتجمع النازل.

3- عدد العمال الذين يتقاضون اجر اكثر من 60 دينار.

4- عدد العمال الذين يتقاضون اجراً اقل من 50 دينار.

الحل :

جدول رقم (10) التكرار المتجمع الصاعد والنازل للاجر الاسبوعي ل(40) عاملاً

الفئات	التكرار (fi)	CF_i^	$CF_i \vee$	Rel - CF_i^	Rel- $CF_i \vee$
30-	3				
40-	1				
50-	8				
60-	10				
70-	7				
80-	7				
90- 100	4				
المجموع	40				

عدد العمال الذين يتقاضون اجراً اكثر من 60 دينار هم

ويمكن حسابها مباشرة من التكرار التجميعي النازل الذي يقابل فئة 60

عدد العمال الذين يتقاضون اجراً اقل من 50 دينار فهم :

5- التمثيل البياني للتوزيعات التكرارية : (الهندسي)

Graphical Presentation Frequency Distribution

يمكن توضيح البيانات بشكل مناسب وذلك باستخدام الرسوم البيانية والصور والاشكال الهندسية بحيث تساعد القارئ على سهولة فهم المعلومات الواردة واستيعاب قيم الظاهرة وبالتالي يمكن مقارنتها مع بعضها ويمكن استخدام مختلف الرسومات البيانية في علم الاحصاء بحيث يمكن رسم البيانات المعروضة بشكل توزيع تكراري اما لعملية الرسم نستخدم المحاور Coordinates المحور السيني X- axis والذي يمثل الفئات ، اما المحور الصادي -Y owi والذي يمثل التكرارات وهناك انواع مختلفة من الرسومات البيانية سوف نتطرق الى بعض منها ، وقبل الدخول بهذا الموضوع يجب الاشارة الى بعض المفاهيم :

1 – البيانات الغير مبوبة : وهي البيانات الأولية او الأصلية التي جمعت ولم تبوب .

مثال : البيانات التالية تمثل اوزان (10) طالب :

64 70 67 62 58 59 56 52 81 72

2– البيانات المبوبة : وهي البيانات التي بوبت ونظمت في جدول توزيع تكراري

مثال : الجدول التالي يمثل توزيع الطلبة حسب أوزانهم

عدد الطلبة (التكرار)	الفئات /الوزن
5	60-62
15	63-65
45	66-68
27	69-71
8	72-74

اذاً التمثيل البياني او الهندسي للتوزيعات التكرارية تكون في حال اذا كانت البيانات :

1– البيانات الغير مبوبة : وهي

أ– الأشرطة البيانية او الأعمدة البيانية Bar Chart :

الرسم هنا عبارة عن مستطيلات راسية تتخذ قواعدها على المحور الافقي لتمثيل الظاهرة بينما ارتفاعها يمثل التكرارات لكل مفردة داخل الظاهرة بينما الارتفاع يمثل تكرار كل مفردة داخل الظاهرة ولرسم الأشرطة البيانية نحتاج الى الخطوات التالية :

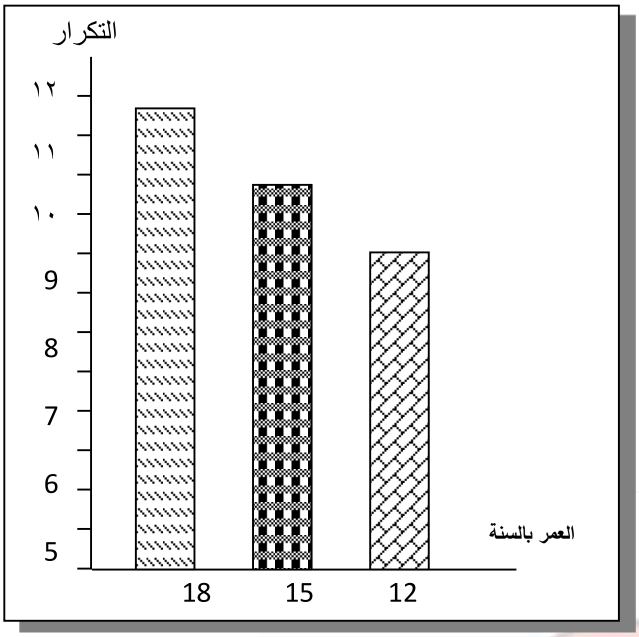
- ندرج المحور الأفقي الى أقسام متساوية بمقياس رسم مناسب بحيث يشمل جميع البيانات فان قاعدة المستطيل تمثل صفة الظاهرة الوصفية ويفضل ترك مسافة متساوية بين المستطيل والاخر .
- اما محور العمودي فيمثل التكرارات ايضاً مقسمة بمقياس رسم مناسب بحيث كل صفة تقابل التكرارات المناسب لها.

وهذه الأعمدة تكون على نوعين :

– **الأشـرطة او الأعمدة البيانية المنفردة** : عبارة عن مجموعة من المستطيلات الراسية او الأفقية تخص صنف واحد فقط للبيانات .

مثال(10)

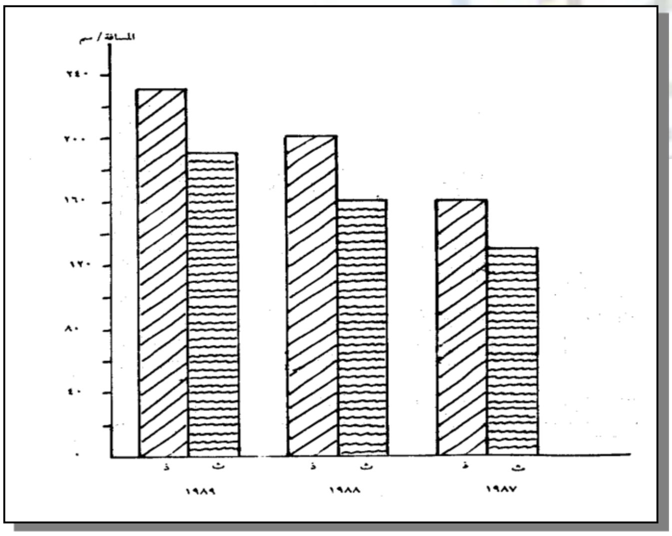
كما في الشكل الذي يوضح السحب على العقلة خلال ثلاث فترات زمنية 1987 ، 1988 و 1989 ، ويمكن ان يمثل هذا الشكل ثلاث مراحل عمرية بدلاً من السنوات مثلا اكثر من 18 سنة واقل من 15 سنة أو اقل من 12 سنة وترسم الاعمدة حسب القيم او البيانات التي يتم الحصول عليها ، وترسم هذه الاعمدة عندما يكون هناك مجالاً كافياً لاستيعاب كافة الاعمدة البيانية



– الأشرطة او الأعمدة البيانية المركبة : وهي اشرطة بيانية تخص صنفين او اكثر للبيانات فمثلاً عدد الطلبة المقبولية في الجامعات للسنوات (1970–1988) مصنفين حسب الجنس ذكور واناث او الحالة الثانية عدد سكان العراق مصنفين حسب الحالة الاجتماعية.

مثال (11)

المقارنة بين اكثر من متغير واحد في الوقت نفسه ، ولكن من المهم في مثل هذه الأشكال ان لا نضع أعمدة اكثر من طاقة الشكل لتجنب الارتباك في المقارنة ويجب تظليل الأعمدة الخاصة بكل متغير أو تمييزها بعلامات تختلف عما في العمود الآخر أو بألوان مختلفة منعا للالتباس ، ويجب كذلك وضع دليل صغير في أعلى الشكل لتبيان كل نوع من أنواع الاعمدة البيانية التي نقارنها . (39: 120-121)



والشكل هذا يوضح مسافة القفز الطويل من الثبات هذا النوع من الرسوم البيانية يمكننا من للذكور والإناث خلال السنوات 1987 ، 1988 ، 1989

ب- المستطيل البياني : عبارة هندسي يستخدم في تمثيل بيانات ظاهرة معينة يمكن تجزئتها الى عدد من الاصناف المقابلة للتجميع مثل : عدد الطلبة موزعين حسب المراحل الدراسية ، ويتم رسم المستطيل البياني باختيار مستطيل ذو قاعدة مناسبة وهذا المستطيل يمثل مجموع البيانات الكلية عن تلك الظاهرة بعد ذلك يتم تجزئة هذا المستطيل الى مستطيلات صغيرة كل واحد منها يمثل جزء من المجموع الكلي للبيانات بحيث ان مساحة المستطيلات الصغيرة تمثل مساحة المستطيل الكبير وتتم عملة لاختيار قواعد المستطيلات الصغيرة كما يلي :

$$\frac{\text{طول قاعدة المستطيل الصغير}}{\text{مجموع البيانات الكلية}} \times \text{طول قاعدة المستطيل الكبير}$$

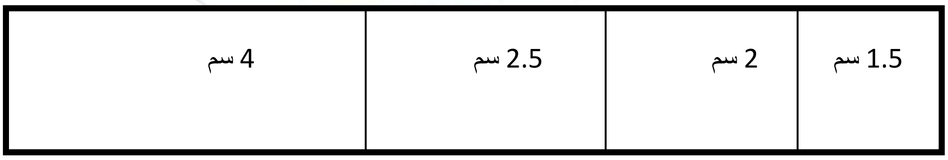
مثال (12) : بلغ عدد الطلبة في احدى الكليات (2000) طالب وطالبة منهم (800) في الصف الاول و(500) في الصف الثاني و (400) في الصف الثالث و(300) في الصف الرابع المطلوب تمثيل هذه البيانات بمستطيل بياني . نختار قاعدة مستطيل مناسبة ولتكن (10)سم فان طول قاعدة كل مستطيل جزئي (صغيره) الخاص بصف معين سيكون كالأتي :

طول قاعدة المستطيل للصف الاول = $\frac{800}{2000} \times 10 = 4$ سم

طول قاعدة المستطيل للصف الثاني = $\frac{500}{2000} \times 10 = 2.5$ سم

طول قاعدة المستطيل للصف الثالث = $\frac{400}{2000} \times 10 = 2$ سم

طول قاعدة المستطيل للصف الرابع = $\frac{300}{2000} \times 10 = 1.5$ سم



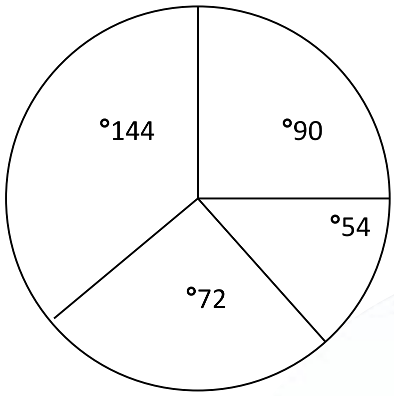
10 سم

ج- الدائرة البيانية : عبارة عن شكل هندسي يكون تمثيل البيانات يشبه المستطيل البياني الا انه بدلاً من المستطيل الكبير تكون دائرة بمساحة معينة وتقسم هذه الدائرة الى قطاعات كل قطاع يمثل صنف من البيانات بحيث ان مجموع مساحات القطاعات يمثل مساحة الدائرة ويتم تحديد كل قطاع بتحديد زاوية كل منهم وكما يلي :

زاوية القطاع = $\frac{\text{عدد بيانات الصف}}{\text{مجموع البيانات الكلية}} \times 360$

بحيث ان مجموع الزوايا لكل القطاعات يساوي 360

مثال (13) في المثال السابق المطلوب تمثيل عدد الطلبة في الكلية بدائرة بيانية.



زاوية القطاع للصف الأول = $\frac{800}{2000} \times 360 = 144°$

زاوية القطاع للصف الثاني = $\frac{500}{2000} \times 360 = 90°$

زاوية القطاع للصف الثالث = $\frac{400}{2000} \times 360 = 72°$

زاوية القطاع للصف الرابع = $\frac{300}{2000} \times 360 = 54°$

مثال : واجب

مثال : بلغ عدد المشـاركين في بطولة الجامعة لهذا العام (200) طالب وطالبة منهم وكان موزعين الكليات حيث بلغ عدد المشـاركين من كلية التربية الرياضية (80) مشارك و(50) مشارك من كلية التربية الأساسية و (40) مشارك من كلية الإدارة والاقتصاد و(30) مشارك من كلية التربية المطلوب تمثيل هذه البيانات بمستطيل بياني ودائرة بيانية.

د- الخط البياني : عبارة عن شكل بياني يوضح التغيرات الحاصلة في ظاهرة معينة خلال فترة معينة من الزمن وهذا الخط يستخدم عادة في حال اجراء مقارنة بين ظاهرتين او اكثر مقاسه بنفس المقياس ويتم رسم الخط البياني كما يلي :

١- رسم المحور الأفقي والمحور العمودي .

٢- يقسم المحور الأفقي الى أجزاء بمقياس رسم مناسب ويمثل هذا المحور وحدات الزمن .

٣- يقسم المحور العمودي الى أجزاء متساوية تمثل وحداة الظاهرة التي يراد توضيحها.

٤- وضع نقطعه امام كل وحدة زمنية ارتفاعها يعادل مقدار الكمية الخاصة بتلك الوحدة.

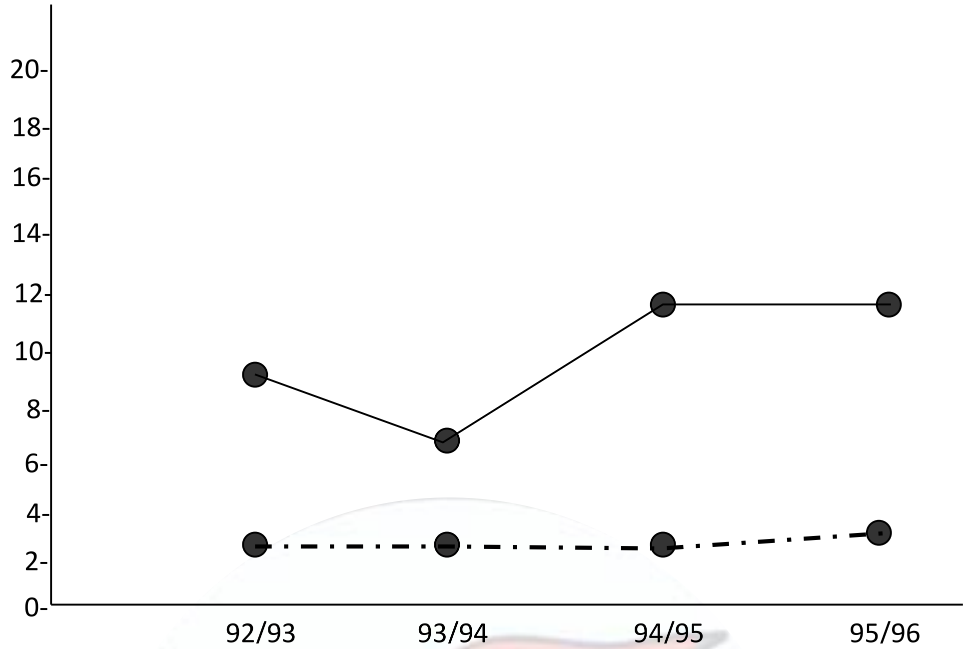
٥- نصل بين نقاط كل ظاهرة بخطوط متصلة وعند وجود اكثر من ظاهرة والتمييز بين الخطوط البيانية برسم كل خط بلون او ان كل خط ياخذ شكل معين مثل

(ــــــــــ) او (..................) او (ــــ ....ــــــ......ــــ) وهكذا.

مثال (14) : الجدول التالي يمثل نشاطات طلبة جامعة الموصل للفترة 93/92 ولغاية 96/95

المطلوب : تمثيل هذه البيانات بخطوط بيانية.

السنوات / النشاطات	93/92	94/93	95/94	96/95
رياضية	9	7	12	12
اجتماعية	2	2	2	3
ثقافية	10	9	9	19
اخرى	6	9	5	8



أكمل الحل :

2- البيانات المبوبة : وهي

أ – المدرج التكراري : عبارة عن مجموعة من المســتطيلات قاعدة كل منها تمثل طول الفئة في التوزيع التكراري وارتفاع كل منها يمثل التكرار المقابل لتلك الفئة وهذه المستطيلات تكون منفصلة عن بعضها في حالة المتغيرات المتقطعة وتكون متصلة في حالة المتغيرات المستمرة وحسب تسلسل فئات التوزيع ولرسم المدرج نتبع الخطوات التالية :

– رسم المحورين الأفقي والعمودي.

– يقسم : المحور الأفقي الى اجزاء بمقياس رسم مناسب ويشكل هذا المحور فئات الجدول ويفضل ترك مسافة صغيرة بين نقطة الصفر والحد الأدنى للفئة الأولى اذا كانت بداية الفئة الأولى لاتساوي صفر.

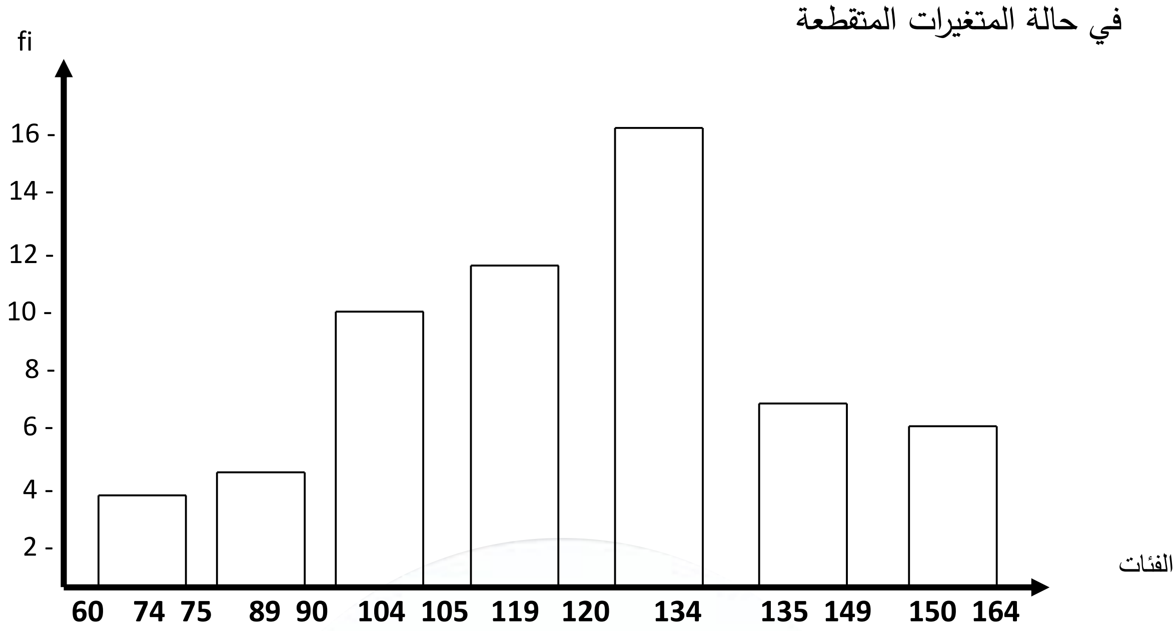
– يقسم المحور العمودي الى أجزاء متساوية تمثل التكرارات .

– يرسم على كل فئة مستطيلات راسياً قاعدته تمثل طول تلك الفئة وارتفاعه يمثل تكرار الفئة ومجموعة هذه المستطيلات تسمى بالمدرج التكراري.

مثال (15) : من الجدول التالي ارسم المدرج التكراري

f_i	الفئات	عدد الفئات
4	60 — 74	1
5	75 — 89	2
10	90 — 104	3
12	105 — 119	4
16	120 — 134	5
7	135 — 149	6
6	150 — 164	7

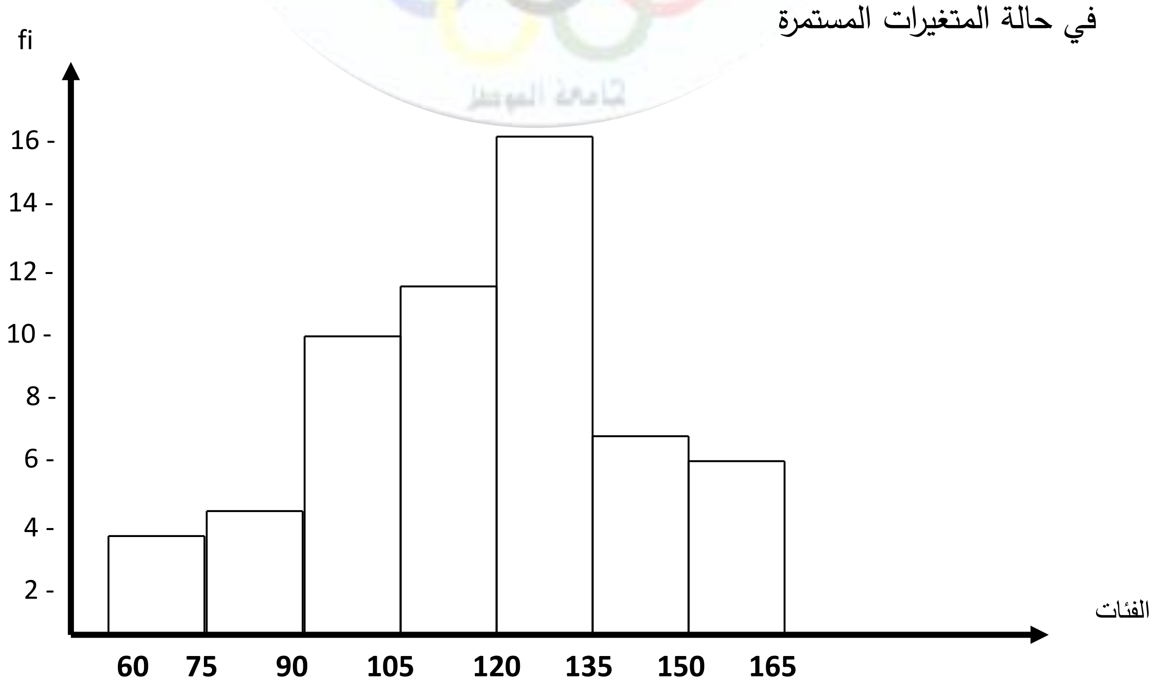
في حالة المتغيرات المتقطعة



مثال (16) : من الجدول التالي ارسم المدرج التكراري

عدد الفئات	الفئات	f_i
1	60 — 75	4
2	75 — 90	5
3	90 — 105	10
4	105 — 120	12
5	120 — 135	16
6	135 — 150	7
7	150 — 165	6

في حالة المتغيرات المستمرة



ب- المضلع التكراري : عبارة عن عدد من المستقيمات المتصلة مع بعضها على شكل سلسلة ونقطة اتصال المستقيم بالأخر تقابل مراكز الفئة وخطوات رسم المضلع كما يلي :

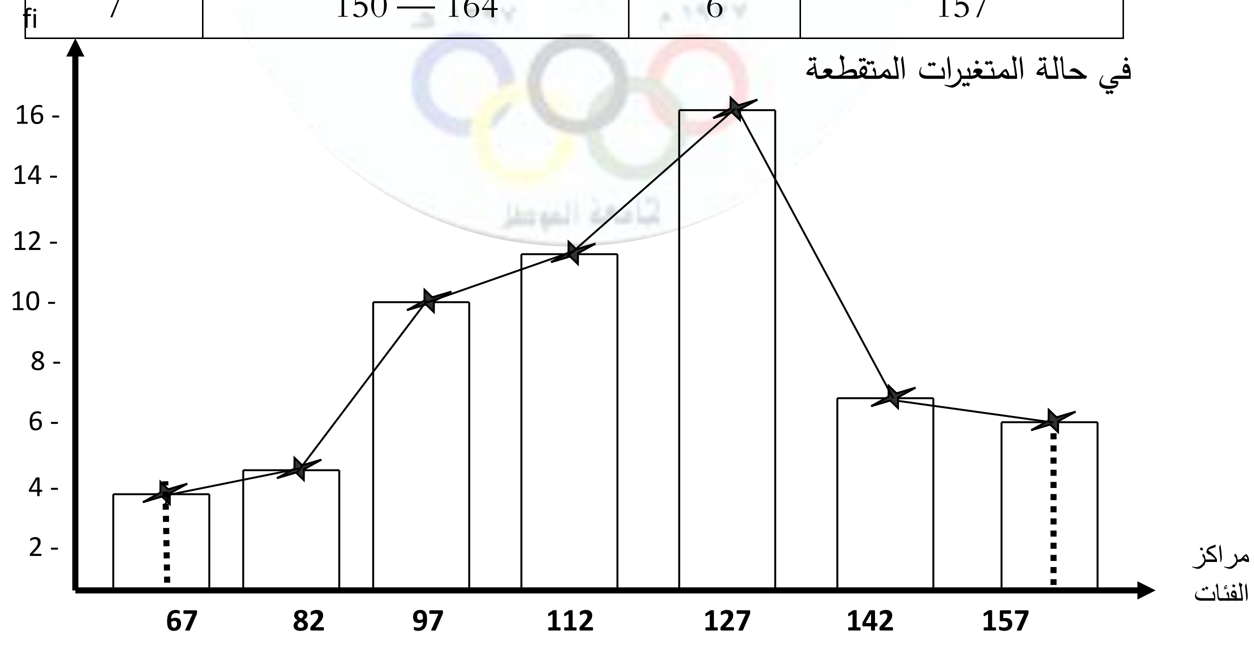
- تمثيل مراكز الفئات على المحور الأفقي.
- تمثيل التكرارات على المحور العمودي.
- تعيين نقطة امام مراكز كل فئة على ارتفاع يعادل تكرار تلك الفئة.
- وصل خطوط مستقيمة بين النقاط المعينة وهي الخطوط تمثل المضلع التكراري.
- يفضل غلق المضلع التكراري مع المحور الأفقي وذلك باختيار مراكز فئة وهي قبل مراكز الفئة الأولى وآخر بعد مركز الفئة الأخيرة ونفترض ان تكرار هذين المركزين مساوي للصفر وتتم عملية الغلق بخط فقط.

ملاحظة : كذلك يمكن رسم المضلع التكراري باستعمال المدرج التكراري ويتم ذلك بتنصيف القواعد العليا للمستطيلات (والتي تمثل مراكز الفئات) بنقاط يتم توصيل هذه النقاط بمستقيمات وكذلك يجب تعديل التكرارات في حالة اختلاف اطوال الفئات عند رسم المضلع التكراري.

مثال (17) : من الجدول التالي ارسم المضلع التكراري

عدد الفئات	الفئات	التكرار f_i	مراكز الفئات x_i
1	60 — 74	4	67
2	75 — 89	5	82
3	90 — 104	10	97
4	105 — 119	12	112
5	120 — 134	16	127
6	135 — 149	7	142
7	150 — 164	6	157

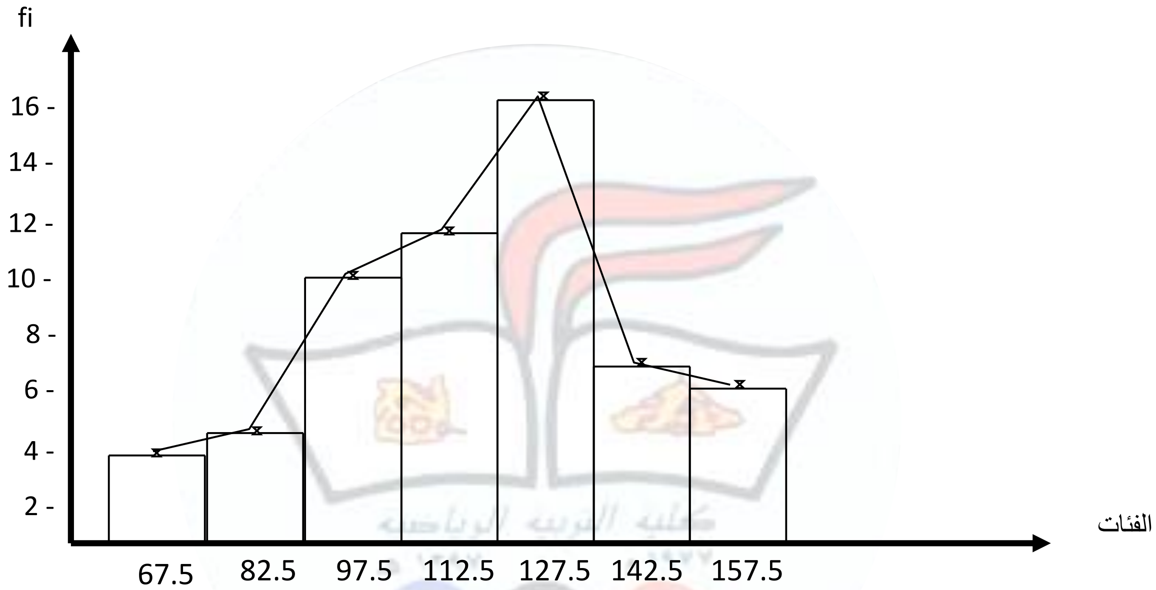
في حالة المتغيرات المتقطعة



مثال (18) : من الجدول التالي ارسم المضلع التكراري

عدد الفئات	الفئات	التكرار f_i	مراكز الفئات x_i
1	60 — 75	4	67.5
2	75 — 90	5	82.5
3	90 — 105	10	97.5
4	105 — 120	12	112.5
5	120 — 135	16	127.5
6	135 — 150	7	142.5
7	150 — 160	6	157.5

في حالة المتغيرات مستمرة



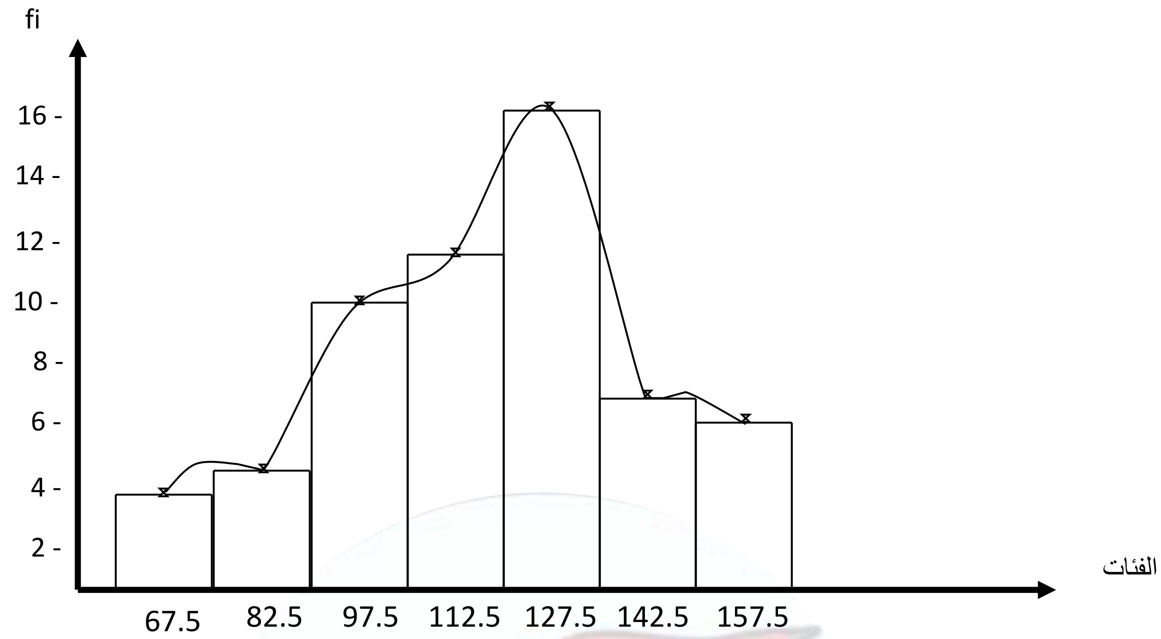
ج- المنحني التكراري : لا تختلف فكرة رسم المنحني التكراري عن المضلع التكراري من حيث الأسلوب لكن الفرق الوحيد بينهما انه بدلاً من توصيل النقاط بمستقيمات نمرر منحني من بين هذه النقاط وكذلك يمكن رسم المنحني التكراري من خلال المدرج التكراري.

ملاحظة : ان المنحني التكراري يتم رسمه للتوزيعات التكراري الخاصة بالمتغيرات من النوع المستمر فقط.

مثال (19) : نفس المثال (18) ارسم المنحني التكراري

عدد الفئات	الفئات	التكرار f_i	مراكز الفئات x_i
1	60 — 75	4	67.5
2	75 — 90	5	82.5
3	90 — 105	10	97.5
4	105 — 120	12	112.5
5	120 — 135	16	127.5
6	135 — 150	7	142.5
7	150 — 160	6	157.5

في حالة المتغيرات مستمرة



د- منحنيات التوزيعات التكرارية المتجمعة :

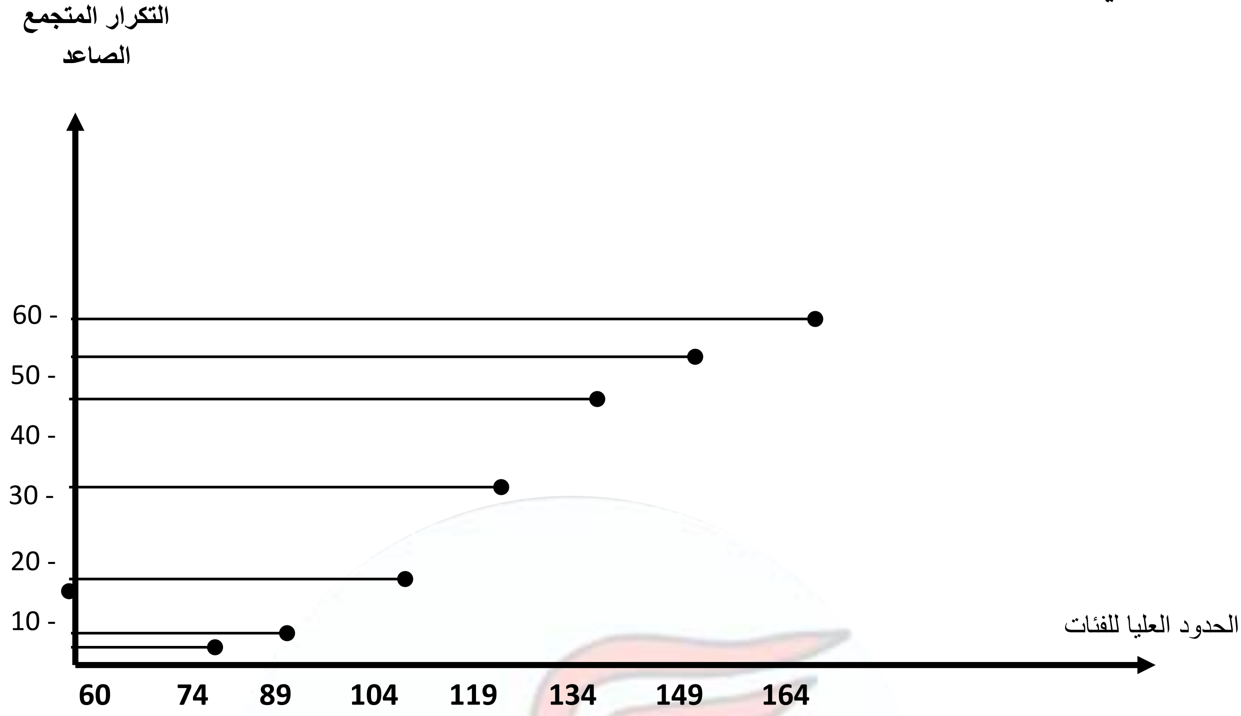
أ- منحني التوزيع التكراري المتجمع الصـــاعد : لغرض تمثيل هذا النوع من التوزيعات يتوجب تحديد نوع المتغير (متقطع – مستمر) ومن ثم تمثيل الحدود العليا للفئات على المحور الافقي والتكرار المتجمع الصاعد على المحور العمودي بعد ذلك يتم تعيين النقاط الخاصة بالتكرارات المتجمعة الصاعدة امام الحدود العليا للفئات.

مثال (20) : الجدول الأتي يبن توزيع تكراري لمجموعة من الطلبة

لمطلوب رسم منحني التوزيع التكراري المتجمع الصاعد

عدد الفئات	الفئات	التكرارf_i	حدود العليا	CF^
1	60 — 74	4	74فاقل	4
2	75 — 89	5	89فاقل	9
3	90 — 104	10	104فاقل	19
4	105 — 119	12	119فاقل	31
5	120 — 134	16	134فاقل	47
6	135 — 149	7	149فاقل	54
7	150 — 165	6	165فاقل	60

في حالة المتغيرات المتقطعة

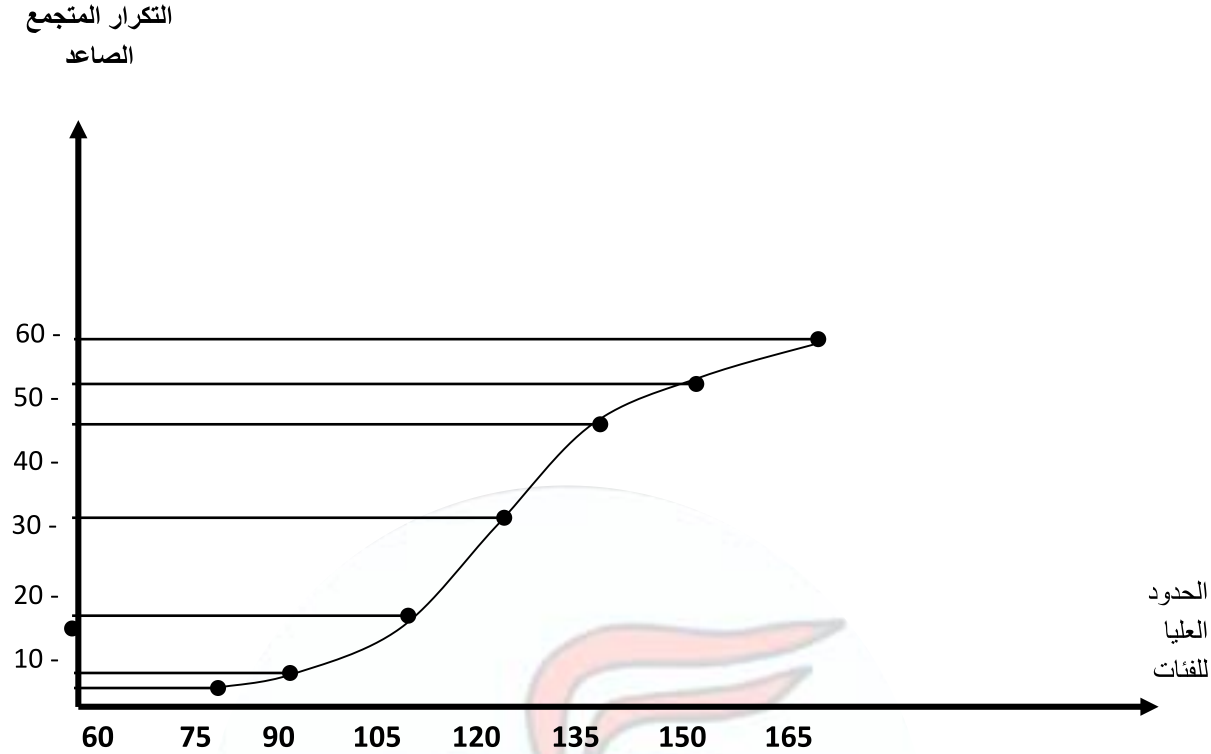


مثال (21) : نفس المثال (18)

المطلوب رسم منحني التوزيع التكراري المتجمع الصاعد ؟

عدد الفئات	الفئات	التكرارf_i	حدود العليا	CF^
1	60 — 75	4	74فاقل	4
2	75 — 90	5	89فاقل	9
3	90 — 105	10	104فاقل	19
4	105 — 120	12	119فاقل	31
5	120 — 135	16	134فاقل	47
6	135 — 150	7	149فاقل	54
7	150 — 165	6	165فاقل	60

في حالة المتغيرات مستمرة



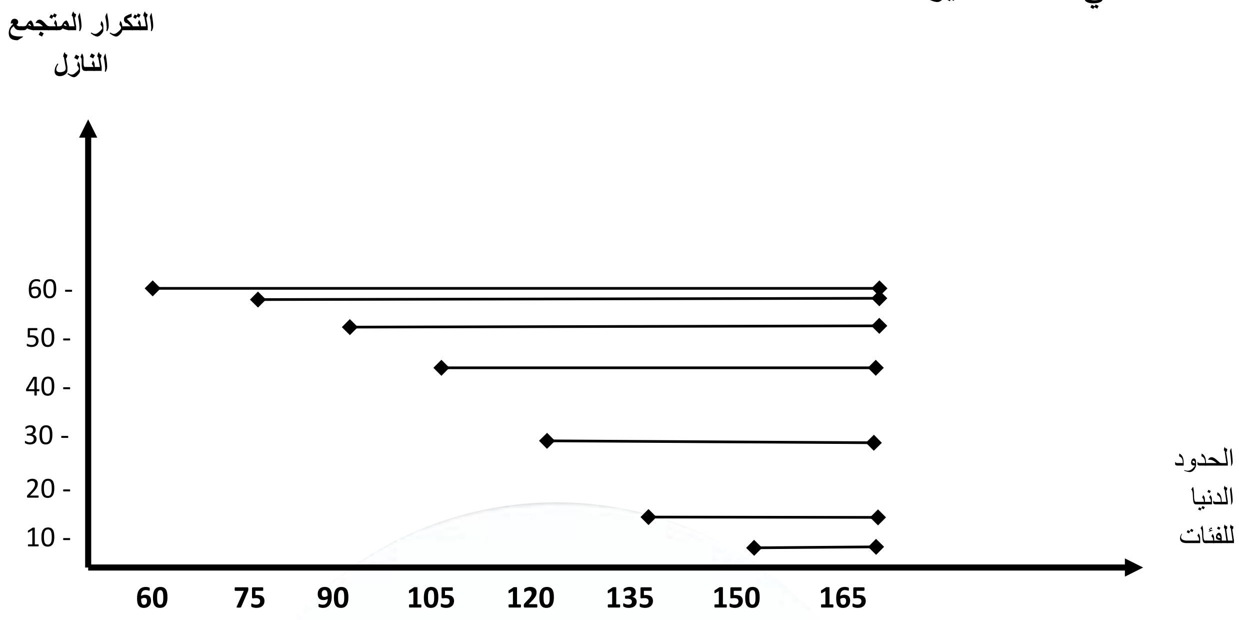
ب– منحني التوزيع التكراري المتجمع النازل : ان طريقة رسم التوزيع المتجمع النازل هي نفس طريقة رسم المتجمع الصاعد عدا انه في حالة التوزيع المتجمع النازل سنمثل المحور العمودي بالتكرار المتجمع النازل بدلاً من التكرار المتجمع الصاعد بعد ذلك يتم تعيين النقاط الخاصة بالتكرارات المتجمعة النازل امام الحدود الدنيا للفئات.

مثال (22) : الجدول الأتي يبن توزيع تكراري لمجموعة من الطلبة.

المطلوب رسم منحني التوزيع التكراري المتجمع النازل ؟

عدد الفئات	الفئات	التكرارf_i	حدود الدنيا	CFv
1	60 — 74	4	60فاكثر	60
2	75 — 89	5	75فاكثر	56
3	90 — 104	10	90فاكثر	51
4	105 — 119	12	105فاكثر	41
5	120 — 134	16	120فاكثر	29
6	135 — 149	7	135فاكثر	13
7	150 — 165	6	150فاكثر	6

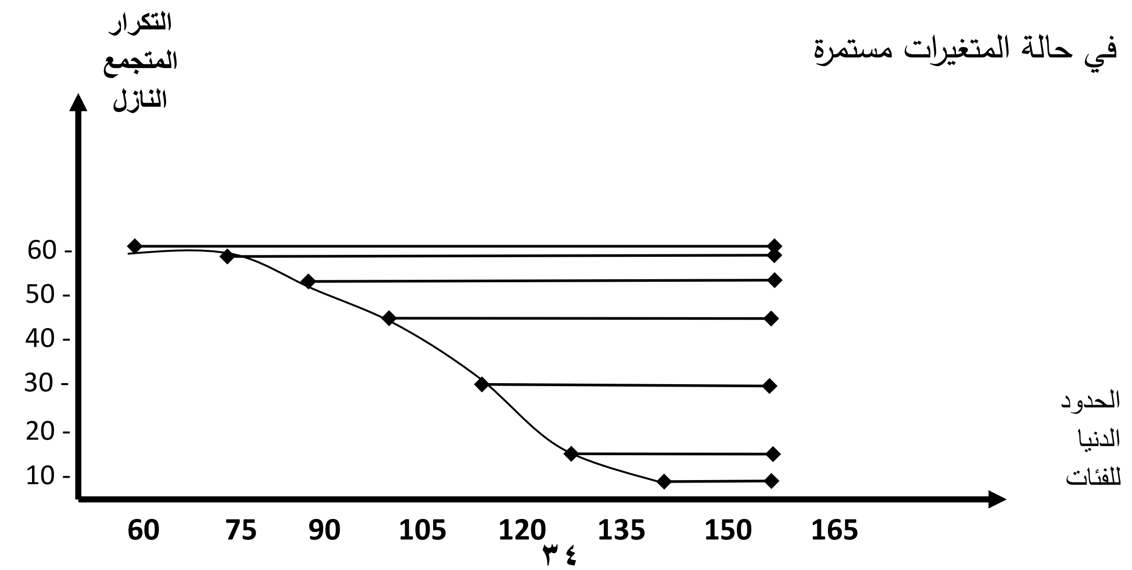
في حالة المتغيرات المتقطعة



مثال (23) : **المطلوب** رسم منحني التوزيع التكراري المتجمع النازل ؟

عدد الفئات	الفئات	التكرار f_i	التكرار المتجمع النازل
1	60 — 75	4	60
2	75 — 90	5	56
3	90 — 105	10	51
4	105 — 120	12	41
5	120 — 135	16	29
6	135 — 150	7	13
7	150 — 160	6	6

في حالة المتغيرات مستمرة



الفصل الثاني

مقاييس النزعة المركزية Measures of Central Tendency

مقدمة :

قبل الشــروع بدراسـة الموضـوعات ذات العلاقة بطرق احتسـاب المؤشـرات الإحصـائية (مقاييس النزعة المركزية) للبيانات المتاحة لا بد من استعراض اهم الرموز والمصطلحات.

1- رمز الجمع ويرمز له بالرمز ($\sum$)

غالباً ما نحتاج وعند التعامل مع الطريقة الإحصـائية في التحليل الى عملية الجمع سلسلة من الأعداد والكميات حيث ان الرمز ($\sum$) يشـير الى عملية الجمع وهو حرف اغريقي يلفظ (Sigmai) وان (i) تمثل دليل لتسـلسـل العدد عند عملية الجمع فإذا كانت (i= 1) ذلك يعني العدد الأول واذا كانت (i=5) فذلك يعني العدد الخامس وهكذا.

وبشكل عام اذا كانت هنالك سلسلة من الكميات عددها (n) فان المجموع الكلي لها يتم التعبير عنه بالشكل ($\sum_{i=1}^{n} x_i$) **وعادة يرمز للدليل بأحد الأحرف الصغيرة مثل (i, j,h)**

خصائص عملية الجمع

على فرض ان (x_i) تمثل سلسلة من البيانات عن المتغير (x) وان (y_i) تمثل عناصر سلسلة من البيانات عن المتغير (y) قوام كل منها (n) من العناصـر وان (a,b) ثابتين حقيقيين فان الخواص ستكون كالأتي :

1- مجموع الكمية الثابتة (a الى n) من المرات يمثل حاصل ضرب (n) بقيمة الثابت (a) أي ان :

$$\sum_{i=1}^{n} a = na$$

مثال :

$$\sum_{i=1}^{6} 4 = 6 \times 4 = 24$$

2- مجموع حاصل ضرب الثابت (a) بقيم عناصر السلسلة (x_i) يمثل حاصل ضرب الثابت (a) بمجموع قيم عناصر سلسلة

$$\sum_{i=1}^{n} a x_i = a \sum_{i=1}^{n} x_i$$

<u>مثال :</u> ليكن

$x_i = \quad 2 \quad , \quad 1 \quad , \quad 3$

فان

$$\sum_{i=1}^{n} 2x_i = 2\sum_{i=1}^{3} x_i = 2(x1+x2+x3) = 2(2+1+3) = 12$$

3- مجموع حاصل اضافة او طرح الثابت (a) الى او من عناصر السلسلة (x_i) يمثل مجموع عناصر السلسلة (x_i) مضافاً له او مطروحاً من حاصل ضرب الثابت (a) بعدد العناصر (n).

$$\sum_{i=1}^{n} (x_i \pm a) = \sum_{i=1}^{n} x_i \pm na$$

<u>مثال :</u> ليكن a=1

$x_i = \quad 2 \quad , \quad 1 \quad , \quad 3$

فان

$$\sum_{i=1}^{n} (x_i - 1) = \sum_{i=1}^{3} x_i - n(1) = (x1+x2+x3) - n(1) = 6 - 3(1) = 3$$

4- مجموع حاصل جمع او طرح بين عناصر السلسلة (x_i) وعناصر السلسلة (y_i) المقابلة له يمثل حاصل جمع مجموع عناصر السلسلتين او الفرق بين مجموع عناصر السلسلتين :

$$\sum_{i=1}^{n} (x_i \pm y_i) = \sum_{i=1}^{n} x_i \pm \sum_{i=1}^{n} y_i$$

او ان :

$$\sum_{i=1}^{n} (ax_i \pm by_i) = a\sum_{i=1}^{n} x_i \pm b\sum_{i=1}^{n} y_i$$

<u>مثال : ليكن</u> a=3 b=2

وان

$x_i = \quad 3 \quad , \quad 2 \quad , y_i = \quad 4 \quad , \quad 3$

فان :

$$\sum_{i=1}^{2} (x_i + y_i) = a\sum_{i=1}^{2} x_i + b\sum_{i=1}^{2} y_i$$

$$3(3+2) + 2(4+3)$$

$$3(5) + 2(7) = 15 + 14 = 29$$

5- ان مجموع عناصر السلسلة (x_i) يمكن تجزئته الى حاصل جمع مجموع مجموعتين جزئيين او اكثر:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{k} x_i + \sum_{i=K+1}^{n} x_i$$

6- مجموع حاصل ضرب عناصر السلسلة (x_i) بعناصر السلسلة (y_i) المقابل لها سيكون كالاتي :

$$\sum_{i=1}^{n}(x_i y_i) = x_1 y_1 + x_2 y_2 + x_3 y_3 + \dots\dots\dots\dots + x_n y_n$$

مثال :

$x_i =$ 2 , 1 , 3 , 5 $yi_i =$ 3 , 2 , 1 , 4 ،

فان :

$$\sum_{i=1}^{n}(x_i y_i) = x_1 y_1 + x_2 y_2 + x_3 y_3 + \dots\dots\dots\dots + x_n y_n$$

$$\sum_{i=1}^{4}(x_i y_i) = (2)(3) + (1)(2) + (3)(1) + (5)(4) = 31$$

$$\sum_{i=1}^{n} \frac{1}{x_i} = \frac{1}{x1} + \frac{1}{x2} + - - - - - + \frac{1}{xn}$$

$$(\sum_{i=1}^{n} yi)^2 = (y1 + y2 + - - - yn)^2$$

أو

$$\sum_{i=1}^{n} yi^2 = (y^2 1 + y^2 2 + - - - y^2 n)$$

واجب : ليكن

$x_i =$ 2 , 4 , 5 , 8, 3

$y_i =$ 1 , 2 , 2 , 3, 3

اوجد ما يأتي :

$1 - \sum_{i=1}^{n} x_i$ $2 - \sum_{i=1}^{n} x_i^2$ $3 - \sum_{i=1}^{n} 3x_i$ $4 - \sum_{i=1}^{n} (x_i - 4)$

$5 - (\sum_{i=1}^{n} y_i)^2$ $6 - \sum_{i=1}^{n} \frac{x_i}{2}$ $7 - \sum_{i=1}^{n} \frac{1}{(y_i)^2}$ $8 - \sum_{i=1}^{n} x_i y_i$

مقاييس النزعة المركزية Measures of Central Tendency

درسنا فيما سبق اهم اساليب جمع وتصنيف البيانات وكيفية تمثيلها في جداول ورسوم هندسية وبيانية وفي هذا الفصل سوف ندرس كيفية تمثيل مجموعة من البيانات بقيمة واحدة فقط من خلال قياس يسمى بمقياس النزع المركزية او مقياس التوسط والمقصود به هو ذلك المقياس الذي يبحث في تقدير قيمة تتمركز حولها أغلبية البيانات الخاصة بظاهرة معينة وهناك عدة مقاييس النزعة المركزية اهمها :

١ - الوسط الحسابي Arithmatic mean

٢ - الوسط الحسابي المرجح Weighted mean

٣ - المنوال mode

٤ - الوسيط median

١ - الوسط الحسابى (المتوسط): ويسمى في بعض الأحيان الوسط او المتوسط او المعدل الحسابي ويرمز له $\bar{X}$ ، يعتبر من أهم مقاييس النزعة المركزية والأكثر استخداما فى الإحصاء والحياة العملية، إذ يستخدم عادة فى الكثير من المقارنات بين الظواهر المختلفة .ولو أسندت قيمة المتوسط لكل مشاهدة فإن مجموع هذه القيم الجديدة يكون مساويا لمجموع المشاهدات الأصلية

طرق احتساب الوسط الحسابي:

أ- في حالة البيانات غير المبوبة

لتكن $(x_1, x_2, x_3 ....x_n)$ تمثل قيم المتغير (x) عددها (n) فان الوسط الحسابي لهذه القيم يحسب كالاتي :

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}[x_1 + x_2 + ... + x_n]$$

حيث ان : $\bar{X}$ يمثل الوسط احسابي $\sum_{i=1}^{n} x_i$ يمثل مجموع القيم او المشاهدات n يمثل عدد القيم

مثال : البيانات التالية تمثل اوزان عينة من الطلبة عددها (5) طالب احسب متوسط وزن الطالب في العينة.

50.2 , 60 , 68.3 , 59.5 , 58.2

$$\bar{X} = \frac{\sum_{i}^{n} x_i}{n}$$

$$\bar{X} = \frac{50.2 + 60 + 68.3 + 59.5 + 58.2}{5}$$

$$\bar{X} = \frac{296.2}{5} = 59.24$$

ب- في حالة البيانات المبوبة : - في حالة التكرارات فقط :

$$\bar{X} = \frac{\sum_{i}^{n} x_i f_i}{\sum_{i}^{n} f_i}$$

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots\dots\dots + x_n f_n}{f_1 + f_2 + \dots\dots\dots + f_n}$$

حيث ان x_i يمثل قيمة المشاهدة f_i تمثل التكرار لكل مشاهدة

مثال : أحد المدربين قام بإجراء اختبار لـــــ (21) لاعباً في اختبار السحب على العقلة فحصلوا على التكرارات المبينة في الجدول الاتي وأراد المدرب ان يحسب الوسط الحسابي لأداء اللاعبين

x_i	f_i **عدد المحاولات**	$x_i f_i$
4	**1**	**4**
5	**4**	**20**
6	**3**	**18**
7	**2**	**14**
8	**5**	**40**
10	**6**	**60**
	$\sum_{i}^{n} f_i = 21$	$\sum_{i}^{n} x_i f_i = 156$

$$\bar{X} = \frac{\sum_{i}^{n} x_i f_i}{\sum_{i}^{n} f_i}$$

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots\dots\dots + x_n f_n}{f_1 + f_2 + \dots\dots\dots + f_n} = \frac{156}{21} = 7.43$$

- **في حالة الفئات والتكرارات :** في هذه الحالة تكون البيانات معطاة في توزيع تكراري يحتوي على m من الفئات مراكزها $(x_1, x_2, \dots\dots\dots, x_m)$ والتكرارات المقابلة لها $(f_1, f_2, \dots\dots\dots, f_m)$ فان الوسط الحسابي لهذا التوزيع يحسب من الصيغة التالية :

$$\bar{X} = \frac{\sum_{i}^{m} x_i f_i}{\sum_{i}^{m} f_i}$$

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots\dots + x_m f_m}{f_1 + f_2 + \dots\dots\dots + f_m}$$

حيث ان x_i يمثل مراكز الفئات و f_i يمثل التكرارات

مثال (٢): احسب متوسط أعمار الطلاب $\bar{x}$ للبيانات التالية:

فئات العمر	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل: لسهولة الحل نضع الجدول التالي:

الفئات	مراكز الفئات (x)	التكرار (f)	$x\ f$
5-6	5.5	2	11
7-8	7.5	5	37.5
9-10	9.5	8	76
11-12	11.5	4	46
13-14	13.5	1	13.5
المجموع		20	184

$$\bar{X} = \frac{\sum_{i}^{m} x_i f_i}{\sum_{i}^{m} f_i}$$

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots\dots + x_m f_m}{f_1 + f_2 + \dots\dots\dots + f_m} = \frac{11+37.5+76+46+13.5}{20} = \frac{184}{20} = 9.2$$

مثال : الاتي توزيع تكراري لدرجات الحرارة في مدينة معينة مســـجلة لمدة (95) يوم متتالي ، المطلوب حساب متوسط درجة الحرارة في هذه المدينة خلال تلك الفترة ؟

درجة الحرارة	عدد الايام f_i	مركز الفئة x_i	$x_i f_i$
0-	4	0.5	2
1-	8	1.5	12
2-	12	2.5	30
3-	16	3.5	56
4-	20	4.5	90
5-	25	5.5	137.5
6-	6	6.5	39
7-8	4	7.5	30
	$\sum_{i}^{m} f_i = 95$		$\sum_{i}^{m} x_i f_i = 396.5$

$$\overline{X} = \frac{396.5}{95} = 4.174^0$$

<u>مثال :</u> تم اختبار طلبة الصـــــفوف الرابعة في مادة الإحصـــــاء البالغ عددهم (74) طالبا وقد تراوحت درجاتهم بين (5) درجات كحد أدنى و (47) كحد أعلى ثم قام مدرس المادة بعمل توزيع تكراري كما في الجدول ادناه مع العلم ان طول الفئة (5) .

درجات الامتحان	عدد الطلاب f_i	مركز الفئة x_i	$x_i f_i$
5-9	2	7	14
10-14	4	12	48
15-19	3	17	51
20-24	2	22	44
25-29	25	27	675
30-34	11	32	352
35-39	14	37	518
40-44	8	42	336
45-49	5	47	235
	$\sum_i^m f_i = 74$		$\sum_i^m x_i f_i = 2273$

$$\overline{X} = \frac{\sum_i^m x_i f_i}{\sum_i^m f_i}$$

$$\overline{X} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots\dots\dots + x_m f_m}{f_1 + f_2 + \dots\dots\dots\dots + f_m} = \frac{2273}{74} = 30.72$$

- بعض مميزات الوسط الحسابي

١ - مقياس سهل حسابه ويخضع للعمليات الجبرية بسهولة.

٢ - يأخذ فى الاعتبار جميع القيم محل الدراسة.

٣-أكثر المقاييس استخداماً فى الإحصاء

- بعض عيوب الوسط الحسابي:

١ - يتأثر بالقيم الشاذة (المتطرفة) وهى القيم الكبيرة جدا أو الصغيرة جدا مقارنة ببقية القيم.

٢ - يصعب حسابه فى حالة الجداول التكرارية المفتوحة، حيث يتطلب ذلك معرفة مركز كل فئة.

٣ - لا يمكن حسابه فى حالة البيانات الوصفية.

٢- الوسط الحسابي المرجح Weighted mean

في بعض الأحيان لا يمكن اعتماد قيمة الوسط الحسابي المحسوبة سابقاً حيث تعطى في هذه الحالة كافة مفردات العينة نفس الأهمية وليس هنالك تفضـيل لمفردة على الأخرى

الا انه من الناحية العملية هنالك الكثير من الحالات تكون لبعض المفردات أهمية اكثر من غيرها مما يستوجب اخذ هذه الأهمية بنظر الاعتبار عند حساب الوسط الحسابي، فمثلاً عند حساب عدد الساعات الأسبوعية المخصصة لكل مادة تدخل في حساب المعدل ويرمز لها بالرمز $\bar{X}w$.

طرق حساب الوسط الحسابي المرجح

أ- في حالة البيانات غير المبوبة :

لتكن $(x_1, x_2, x_3 .... x_n)$ تمثل قيم المتغير (x) عددها (n) وان $(w_1, w_2, w_3 .... w_n)$ اوزان لهذه القيم فان الوسط الحسابي المرجح يحسب وفق الصيغة التالية :

$$\bar{X}w = \frac{\sum_i^n w_i x_i}{\sum_i^n w_i}$$

$$\bar{X}w = \frac{w_1 x_1 + w_2 x_2 + ... + w_n x_n}{w_1 + w_2 + ... + w_n}$$

مثال : كانت درجات احد الطلبة كالاتي ، المطلوب : حساب معدل الطالب ؟

الدرجات x_i	عدد الساعات الأسبوعية w_i	$w_i x_i$
62	2	124
80	2	160
75	2	150
88	3	264
84	3	252
86	3	258
90	3	270
	$\sum_i^n w_i = 18$	$\sum_i^n w_i x_i = 1478$

$$\bar{X}w = \frac{\sum_i^n w_i x_i}{\sum_i^n w_i}$$

$$\bar{X}w = \frac{124 + 160 + 150 + 264 + 252 + 258 + 271}{2 + 2 + 2 + 3 + 3 + 3 + 3} = \frac{1478}{18} = 82.111$$

مثال : اذا كان متوسط سن العامل في شركة A هي 35 سنة وشركة B هي 40 وشركة C هي 45 فاذا كان عدد العاملين في الشركات الثلاثة على التوالي :300 ، 200 ،100 فما هو متوسط سن العاملين في الشركات الثلاثة.

$$\bar{X}w = \frac{\sum_{i}^{n} w_i x_i}{\sum_{i}^{n} w_i}$$

$$\bar{X}w = \frac{300\times 35 + 200\times 40 + 100\times 45}{300+200+100} = \frac{23000}{600} = 38.33333$$

ب- في حـــالـــة البيـــانـــات المبوبـــة : لتكن $(x_1, x_2, \ldots\ldots\ldots, x_m)$ تمثـل مراكز الفئـات و $(f_1, f_2, \ldots\ldots, f_m)$ تمثل التكرارات و $(w_1, w_2, w_3 \ldots\ldots w_n)$ اوزان الفئات فان الوســط الحســابي المرجح يحسب وفق الصيغة التالية :

$$\bar{X}w = \frac{\sum_{i}^{n} w_i f_i x_i}{\sum_{i}^{n} w_i f_i}$$

$$\bar{X}w = \frac{w_1 f_1 x_1 + w_2 f_2 x_2 + \ldots\ldots\ldots + w_n f_n x_n}{w_1 f_1 + w_2 f_2 + \ldots\ldots\ldots + w_n f_n}$$

مثال : الاتي توزيع تكراري لانتاج مصنع معين من سلعة معينة لاحد الايام حسب عدد المكائن العاملة في ذلك اليوم وعدد ساعات العمل المحددة لاشتغال كل ماكنة حسب مواصفات المنشأة المطلوب حساب متوسط انتاجية الماكنة الواحدة في هذا المصنع وفي ذلك اليوم ؟

فئات الإنتاج (بالطن)	**عدد المكائن** f_i	**ساعات العمل لمقررة لكل ماكنة** w_i	**مركز الفئة** x_i	$w_i f_i$	$w_i f_i x_i$
2-3	**4**	**6**	**3**	**24**	**72**
4-5	**5**	**5**	**5**	**25**	**125**
6-7	**6**	**6**	**7**	**36**	**252**
8-9	**3**	**4**	**9**	**12**	**108**
10-12	**2**	**4**	**11**	**8**	**88**
				$\sum_{i}^{n} w_i f_i = 105$	$\sum_{i}^{n} w_i f_i x_i = 64$

$$\bar{X}w = \frac{\sum_{i}^{n} w_i f_i x_i}{\sum_{i}^{n} w_i f_i}$$

بـــــالـــــــطـــــن

$$\bar{X}w = \frac{72+725+252+108+88}{24+25+36+12+8} = \frac{645}{105} = 6.142857$$

مثال : من جدول التوزيع التكراري الاتي ، اوجد الوسط الحسابي المرجح ؟

الفئات	f_i	w_i الاوزان	x_i مركز الفئة	$w_i f_i$	$w_i f_i x_i$
10-20	6	1	15	6	90
20-30	10	2	25	20	500
30-40	25	3	35	75	2625
40-50	20	4	45	80	3600
50-60	25	5	55	125	6875
60-70	4	6	65	24	1560
				$\sum_{i}^{n} w_i f_i = 330$	$\sum_{i}^{n} w_i f_i x_i =$ [illegible]50

$$\bar{X}w = \frac{\sum_{i}^{n} w_i f_i x_i}{\sum_{i}^{n} w_i f_i}$$

$$\bar{X}w = \frac{15250}{330} = 46.21212$$

٦- **المنوال mode** : يعرف المنوال انـه تلـك القيمـة التي تتكرر اكثر من غيرها من بين مجموعة من القيم او انها القيمة الشائعة من بين مجموعة من القيم ويرمز للمنوال بـ($M_{\circ}$).

طرق حسـابه : أ- في حالة البيانات الغير مبوبة : لتكن $(x_1, x_2, x_3 ....x_n)$ تمثل قيم المتغير (x) عددها (n) ونفرض ان احد هذه القيم تكررت اكثر من غيرها ولتكن هذه القيمة (x_j) عندئذ وحسب تعريف المنوال فان (x_j) تمثل المنوال لهذه المجموعة.

مثال : احسب المنوال للبيانات التالية :

2, 3, 2, 4, 2, 5, 4, 4, 5, 4, 6, 7, 8

اذا فان المنوال هو : $M_{\circ} = 4$

مثال : احسب المنوال للبيانات التالية :

2, 5, 8, 9, 10 ,7

لا يوجد منوال

مثال : احسب المنوال للبيانات التالية :

2, 4, 2, 5, 5, 7, 6, 8

اذا فان المنوال هو : $M_{\circ} = 5,2$

ب- في حالة البيانات المبوبة :

باستخدام طريقة الفروق لبيرسون وهي كالاتي

١- نجد الفئة المنوالية : وهي الفئة التي تقابل اكبر تكراراً بين الفئات.

٢- نجد الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة ونرمز له بالرمز d_1.

٣- نجد الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة وليكن ونرمز له بالرمز d_2.

٤- نطبق القانون الاتي :

$$M_o = a + \left[\frac{d_1}{d_1 + d_2}\right] \times L$$

حيث ان a: تمثل الحد الادنى للفئة المنوالية. L طول الفئة .

مثال : اوجد قيمة المنوال من الجدول التكراري التالي :

الفئات	f_i
50-60	5
60-70	12
70-80	24
80-90	6
90-100	11
100-110	7

الحل :

١- الفئة المنوالية هي التي تقابل اكبر تكرار والمتمثل بقيمة (24).

٢- قيمة d_1 تساوي $d_1 = 24 - 12 = 12$

٣- قيمة d_2 تساوي $d_2 = 24 - 6 = 18$

٤- نطبق القانون الاتي :

$$M_o = a + \left[\frac{d_1}{d_1 + d_2}\right] \times L$$

$$= 70 + \left[\frac{12}{12 + 18}\right] \times 10 = 74$$

مثال : اوجد قيمة المنوال من الجدول التكراري التالي :

الفئات	f_i
20-25	**3**
25-30	**12**
30-35	**31**
35-40	**10**
40-45	**4**

الحل :

١- الفئة المنوالية هي التي تقابل اكبر تكرار والمتمثل بقيمة (31).

٢- قيمة d_1 تساوي $d_1 = 31 - 12 = 19$

٣- قيمة d_2 تساوي $d_2 = 31 - 10 = 21$

٤- نطبق القانون الاتي :

$$M_o = a + \left[\frac{d_1}{d_1 + d_2}\right] \times L$$

$$= 30 + \left[\frac{19}{19 + 21}\right] \times 5 = 32.37$$

يمكن حساب المنوال حساب المنوال بيانياً :

بطريقة الرسم حيث نتبع الخطوات التالية :

١- نرسم مدرج تكراري

٢- نحدد الفئة المنوالية وهي الفئة التي لها اعلى مستطيل

٣- نصل الزاوية العليا اليمنى للفئة المنوالية مع الزاوية العليا اليمنى للفئة السابقة للفئة المنوالية.

٤- نصل الزاوية العليا اليسرى للفئة المنوالية مع الزاوية العليا اليسرى للفئة المنوالية اللاحقة.

٥- من نقطة تقاطع القطرين ننزل عمود على المحور الأفقي فتعطي المنوال.

مثال : من جدول التوزيع التالي اوجد المنوال حسابياً وبيانياً

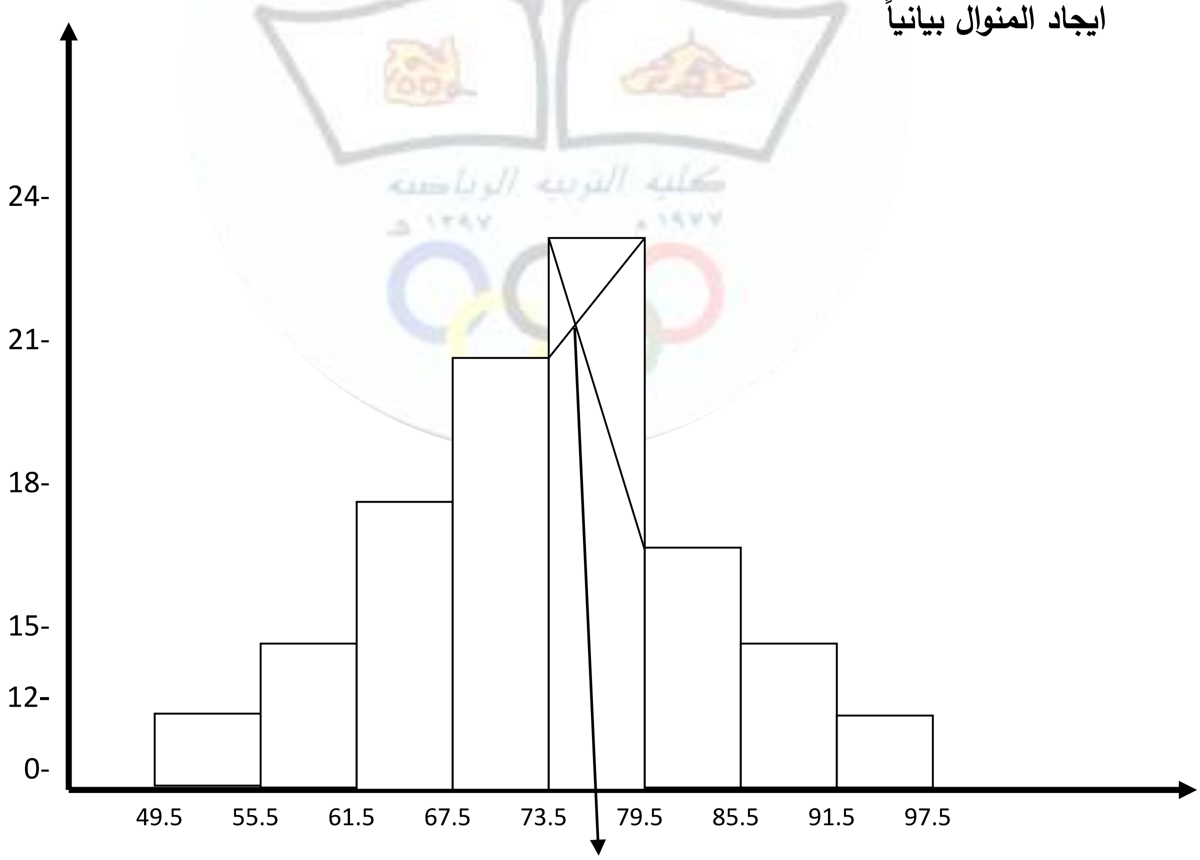
الحدود الفعلية للفئات	f_i	الفئات
49.5-55.5	**1**	**50-55**
55.5-61.5	**5**	**56-61**
61.5-67.5	**12**	**62-67**
67.5-73.5	**15**	**68-73**
73.5-79.5	**22**	**74-79**
79.5-85.5	11	**80-85**
85.5-91.5	6	**86-91**
91.5-97.5	3	**92-97**

d1= 22-15=7
d2= 22-11=11

$$M_o = a + \left[\frac{d_1}{d_1 + d_2}\right] \times L$$

$$= 74 + \left[\frac{7}{7+11}\right] \times 6 = 76.33$$

ايجاد المنوال بيانياً



- **مزايا المنوال :**

١- مقياس سهل الفهم والحساب.

٢- يمكن ايجاد المنوال لبيانات وصـــفية مثلاً : لو كانت تقديرات طالب معين في مجموعة امتحانات هي (متوسط ، متوسط ، مقبول ، جيد ، جيد ، متوسط) فان المنوال في هذه الحالة هو متوسط .

٣- ان المنوال لا يتأثر اطلاقاً بالقيم الشاذة والمتطرفة.

٤- يمكن ايجاد المنوال في حالة التوزيعات التكرارية المفتوحة.

- **عيوب المنوال :**

١- يتأثر وعلى نحو كبير بأخطاء المعاينة.

٢- لا يســـتند عملية إيجاده الى كافة البيانات المتاحة حيث انه بمجرد ملاحظة اكبر تكرار يتم معرفة المنوال او فئته وتهمل كافة القيم الأخرى او الفئات الأخرى

- **الوســيط median** : هو القيمة الوسـطى بين المفردات الاحصــائية ويعرف بانه القيمة التي تقع في منتصف القيم وذلك بعد ترتيبها تصاعديا او تنازليا ويرمز له بالرمز Med

- **في حالة البيانات الغير مبوبة :** عند ترتيب البيانات (المشـاهدات) ترتيبا تصــاعديا أو تنازليا فإن الوسيط يكون البيان (المشـاهدة) التى يقع 50% من البيانات قبلها فى الترتيب و 50 % من البيانات بعدها فى الترتيب .فإذا كان عدد البيانات فرديا فإن الوسـيط يكون المشـاهدة التى تقع فى المنتصـف، وإذا كان عدد البيانات زوجيا فإن الوسـيط هو متوسـط المشـاهدتين اللتين تقعان فى المنتصف.

الفردي=n+1/2

مثال : اوجد الوسيط لدرجات الطلاب حيث كانت درجاتهم كالتالي:

72, 60, 40, 80, 63

الحل :يتم ترتيب البيانات تصاعديا كالتالي5+1/2=3:

40
60
63
72
80

بما أن عدد المشـاهدات فردى فيكون الوسـيط هو المشـاهدة التى تقع فى منتصـف هذه البيانات . وعليه فإن الوسـيط هو المشـاهدة رقم 3 وقيمتها 63 نلاحظ أن هناك بيانين يقعان فى الترتيب قبل الوسيط وبيانين يقعان بعده.

مثال : اوجد الوسيط لدرجات الطلاب الآتية:
72, 60, 72, 40, 80, 63
الحل :يتم ترتيب البيانات تصاعديا كالتالي:
40
60
63
72
72
80

بما أن عدد البيانات زوجى فتكون قيمة الوسـيط هى متوسـط قيمتى المشـاهدتين اللتين يقعان فى المنتصف.

$$\text{Med} = \frac{63+72}{2} = 67.5$$

مثال : جد الوسيط من البيانات التالية :

27, 18, 10, 19, 25, 12

الحل : ١- نرتب البيانات تصاعدياً :

25, 21, 19, 18, 12, 10

٢- عدد البيانات (6) زوجي :

بما أن عدد البيانات زوجى فتكون قيمة الوسـيط هى متوسـط قيمتى المشـاهدتين اللتين يقعان فى المنتصف.

$$\text{Med} = \frac{18+19}{2} = 18.5$$

- في حالة البيانات المبوبة :

خطوات أيجاد الوسيط للبيانات المبوبة

١- أيجاد التكرار المتجمع الصاعد

٢- نجد ترتيب الوسيط باستخدام الصيغة $\frac{\sum_{i}^{n} f_i}{2}$

٣- تحديد الفئة الوسيطية والتي غالباً ما تقابل اكبر تكرار

٤- نطبق القانون التالي :

$$M_e = a + \left[\frac{\frac{\sum_{i}^{n} f_i}{2} - F}{fi} \right] \times W$$

حيث ان $\frac{\sum_{i}^{n} f_i}{2}$ **ترتيب الوسيط**

F التكرار المتجمع الصاعد السابق

f يمثل الفئة الوسيطية والتي غالباً ما تقابل اكبر تكرار

مثال : احسب الوسيط من الجدول التالي :

الفئات	f_i	التكرار المتجمع الصاعد
80-90	8	8
90-100	22	30
100-110	40	70
110-120	19	89
120-130	11	100
المجموع	100	

نجد ترتيب الوسيط $\frac{\sum_{i}^{n} f_i}{2} = \frac{100}{2} = 50$

$$M_e = a + \left[\frac{\frac{\sum_{i}^{n} f_i}{2} - F}{fi}\right] \times W$$

$$= 100 + \left[\frac{50-30}{40}\right] \times 10 = 105$$

مثال : من جدول التوزيع التكراري التالي اوجد :

١- الوسيط حسابياً

٢- الوسيط بيانياً

الفئات	f_i	التكرار المتجمع الصاعد
5-10	2	2
10-15	3	5
15-20	10	15
20-25	8	23
25-30	7	30
المجموع	$\sum_{i}^{n} f_i = 30$	

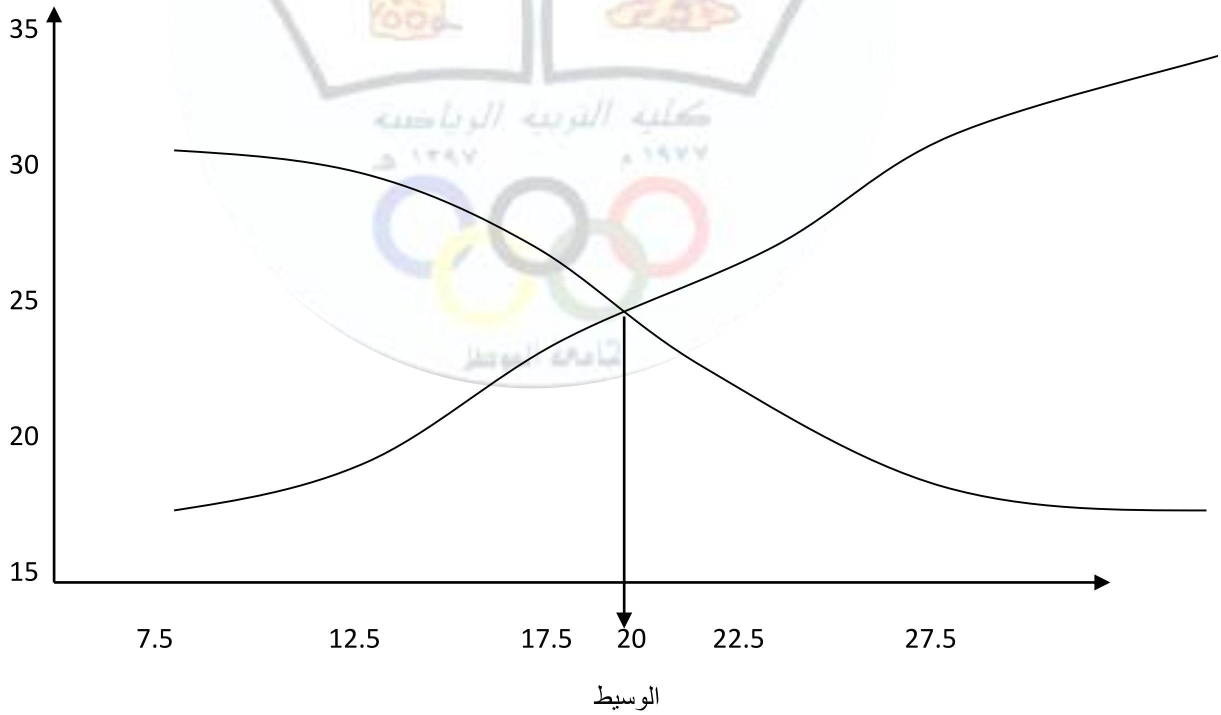
نجد ترتيب الوسيط $\frac{\sum_{i}^{n} f_i}{2} = \frac{30}{2} = 15$

$$M_e = a + \left[\frac{\frac{\sum_i^n f_i}{2} - F}{f} \right] \times W$$

$$= 15 + \left[\frac{15-5}{10} \right] \times 5 = 20$$

٢- لايجاد الوسيط بيانياً نجد التكرار المتجمع الصاعد والنازل ومراكز الفئات

الفئات	f_i	التكرار المتجمع الصاعد	التكرار المتجمع النازل	مراكز الفئات
5-10	**2**	**2**	**30**	**7.5**
10-15	**3**	**5**	**28**	**12.5**
15-20	**10**	**15**	**25**	**17.5**
20-25	**8**	**23**	**15**	**22.5**
25-30	**7**	**30**	**7**	**27.5**
المجموع	$\sum_i^n f_i = 30$			



العلاقة بين الوسط الحسابى والوسيط والمنوال

توجد علاقة تجريبية بين المقاييس الثلاثة (الوسط الحسابى – الوسيط – المنوال) وذلك فى حالة التوزيعات التكرارية أحادية المنوال وغير المتماثلة والمتماثلة وذات الالتواء البسيط .وتعطى هذه العلاقة من خلال المعادلة التالية:

$$\bar{X} - M_0 = 3(\bar{X} - M_e)$$

وقد وجد أن الوسيط تقع قيمته بين قيمتي الوسط الحسابي والمنوال.

وفى حالة التوزيعات التكرارية المتماثلة الوحيدة المنوال فإن قيمة الوسط الحسابي تكون مساوية

لقيمة الوسيط تكون مساوية لقيمة المنوال أى أن:

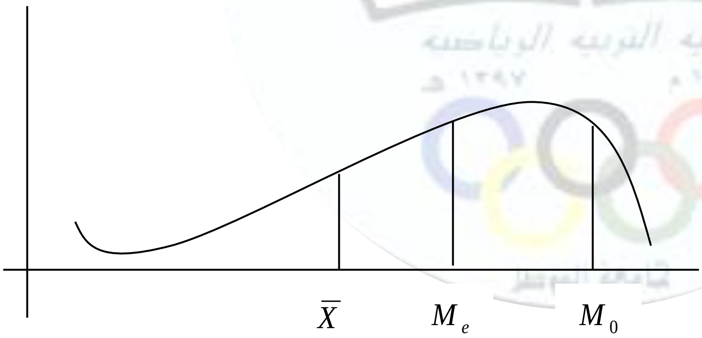
الوسط الحسابي = الوسيط = المنوال

أما فى حالة (التوزيعات التكرارية الملتوية) التواء موجب – التواء سالب فان العلاقة السابقة تكون غير صحيحة

ملاحظات

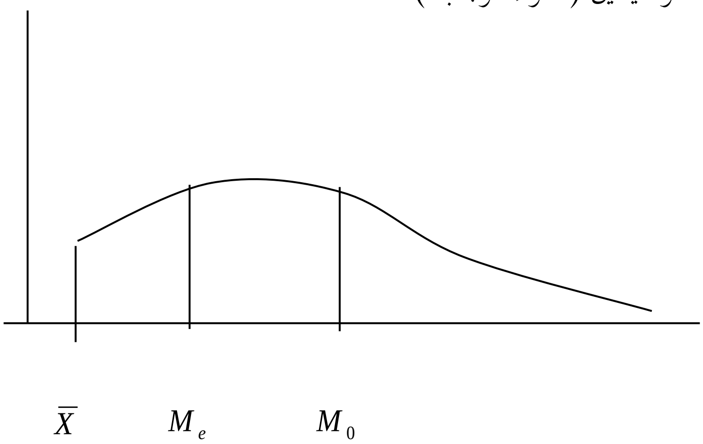
١- يكون المنحني ملتوياً نحو اليسار (اللتواء سالب)

فان $M_0 \geq M_e \geq \bar{X}$



٢- يكون المنحني ملتوياً نحو اليمين (اللتواء موجب)

فان $M_0 < M_e < \bar{X}$



مثال : لديك توزيع تكراري غير متماثل وكان $M_e = 60$ و $\bar{X} = 50$

جد المنوال ونوع الالتواء

الحل /

$$\bar{X} - M_0 = 3(\bar{X} - M_e)$$
$$50 - M_0 = 3(50 - 60)$$
$$M_0 = 80$$

نوع الالتواء (التواء سالب)

$$M_0 \geq M_e \geq \bar{X}$$
$$80 \geq 60 \geq 50$$

مثال : لديك توزيع تكراري غير متماثل وكان $M_0 = 10$ و $M_e = 40$

جد الوسط الحسابي ونوع الالتواء

الحل /

$$\bar{X} - M_0 = 3(\bar{X} - M_e)$$
$$\bar{X} - 10 = 3(\bar{X} - 40)$$
$$\bar{X} = 55$$

نوع الالتواء (التواء موجب)

$$M_0 < M_e < \bar{X}$$
$$10 < 40 < 55$$

الفصل الثالث
مقاييس التشتت

مقدمة : لاحظنا في الفصل السابق كيفية تمثيل مجموعة من البيانات بعدد واحد يعطي فكرة عن تلك المجموعة سميناه " مقياس نزعة مركزية "، ان هذا المقياس لوحدة لايعطي فكرة متكاملة عن مجموعة البيانات وخصوصاً فيما يتعلق الامر بمقدار تجانس مفردات المجموعة مقارنة بمجموعة أخرى او مجموعات أخرى من البيانات عن نفس الظاهرة، لذلك سوف يتم في هذا الفصل استعراض مقاييس أخرى تختص بمقياس درجة تجانس مفردات المجموعة تعطي فكرة اكثر تكاملاً الى جانب مقاييس النزعة المركزية.

وهذه المقاييس تسمى التشتت او الاختلاف ويقصد بالتشتت او الاختلاف بانه التباعد او الانتشار او التقارب الموجود بين قيم مجموعة من المفردات او عن قيمة معينة ثابته (كالوسط الحسابي) مثلاً ، وان لمقاييس التشتت اهميتها في وصف التوزيعات ومقارنتها مع بعضها حيث ان مقاييس التوسط وحدها لاتكتفي لهذا الغرض فمثلاً عند مقارنة المجموعتين الاتيتين :

المجموعة الاول : 9 , 7 , 8 , 9 , 10 , 11

المجموعة الثانية : 3 , 9 , 6 , 9 , 12 , 15

نلاحظ ان الوسط الحسابي لكل من المجموعتين هو (9) ولكن تبدو المجموعة الأولى اكثر تجانساً لذلك فان مقياس التشتت يكون كبيراً علماً كانت البيانات اقل تجانساً ويكون صغيراً عندما يكون الاختلافات بين القيم قليلة .

وهناك عدة مقاييس للتشتت أهما :

أ- مقاييس التشتت المطلقة : وهي تبين درجة التجانس بشكل مطلق وتكون مقاسة بنفس وحدات قياس المتغير ومن أهمها :

١- المدى .

٢- الانحراف المتوسط.

٣- الانحراف المعياري.

٤- التباين .

ب- مقاييس التشتت النسبية : وهي التي تبين درجة التجانس بشكل نسبي وتكون خالية من وحدات القياس ومن أهمها :

١- معامل الاختلاف.

٢- الدرجة المعيارية.

٣- الدرجة المعيارية المعدلة.

أ- مقاييس التشتت المطلقة :

1- المدى : في حالة البيانات الغير مبوبة : ويرمز له بالرمز (R) يعتبر المدى من ابسط أنواع مقاييس التشـــتت المطلقة ويعرف بأنه الفرق بين اكبر قيمة في المجموعة واصـــغر قيما فيها أي ان :

$R= X_U - X_L$ حيث ان

حيث ان X_U اكبر قيمة

X_Lاصغر قيمة

مثال : جد المدى لكل من المجموعتين الآتيتين :

المجموعة الأول : 5 , 18 , 10 , 15 , 3 , 7 , 6 , 12

المجموعة الثانية : 18 9 , 8 , 9 , 8 , 8 , 3 , 9

$R1 =18- 3= 15$

$R2 = 18- 3 = 15$

اما في حالة البيانات المبوبة فيكون المدى في هذه الحالة يمثل الفرق ما بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأول.

ملاحظة : يكون المدى في اكثر من الأحيان مضـــــلاً لأنه يعتمد فقط على القيمتين المتطرفتين ونلاحظ ان المدى في المثال الســـــابق كان لكلاً المجموعتين متســـاوي ولكننا نلاحظ حقيقة ان الاختلاف في المجموعة الأول اكبر من الاختلاف في المجموعة الثانية.

2- الانحراف المتوسـط : يعرف الانحراف المتوسـط بأنه مجموع الانحرافات المطلقة لقيم متغير ما عن وسطه الحسابي مقسوماً على عدد هذه القيم ويرمز له بالرمز (M.D).

حساب الانحراف المتوسط :

أ- البيانـات غير المبوبـة : لتكن $(x_1, x_2, x_3 ....x_n)$ تمثل قيم المتغير (x) عددها (n) وان $(\bar{X})$ يمثل ان الوسـط الحسـابي لهذه القيم فيمكن حسـاب الانحراف المتوسـط حسـب الصـيغة التالية :

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{X}|}{n} =$$

حيث ان : $\bar{X}$ يمثل الوسط الحسابي x_i يمثل القيم المتغير او المشاهدات n يمثل عدد القيم

مثال : احسب الانحراف المتوسط للبيانات التالية :

2 , 3 , 4 , 5 , 5 , 6 , 7 , 10 , 13 , 14 , 19

$$M.D=\frac{\sum_{i=1}^{n}\left|x_i-\bar{X}\right|}{n}$$

$$\bar{X}=\frac{\sum_{i}^{n}x_i}{n}=\frac{2+3+4+5+5+6+7+10+13+14+19}{11}=\frac{88}{11}=8$$

$$M.D=\frac{|2-8|+|3-8|+|4-8|+|5-8|+|5-8|+|6-8|+|7-8|+|10-8|+|13-8|+|14-8|+|19-8|}{11}$$

$$M.D=\frac{48}{11}=4.364$$

ب- في حالة البيانات المبوبة : يمكن حساب الانحراف المتوسط لتوزيع تكراري حسب الصيغة التالية :

$$M.D=\frac{\sum_{i=1}^{n}f_i\quad\left|x_i-\bar{X}\right|}{\sum_{i}^{n}f_{i}}$$

حيث ان f_i يمثل تكرارات الفئة

وان x_i مراكز الفئة

وان $\bar{X}$ يمثل الوسط الحسابي

مثال : الأتي توزيع تكراري لدرجات مجموعة من الطلبة، المطلوب قياس مقدار التشـــتت في درجات هذه المجموعة من الطلبة باستخدام الانحراف المتوسط .

الفئات	f_i	x_i	$f_i x_i$	$x_i-\bar{X}$	$\left\|x_i-\bar{X}\right\|$	$f_i\left\|x_i-\bar{X}\right\|$
0_10	4	5	20	-54	54	216
10_20	4	15	60	-44	44	176
20_30	8	25	200	-34	34	272
30_40	16	35	560	-24	24	384
40_50	25	45	1125	-14	14	350
50_60	58	55	3190	-4	4	232
60_70	42	65	2730	6	6	252
70_80	35	75	2625	16	16	560
80_90	18	85	1530	26	26	468
90_100	10	95	950	36	36	360
المجموع	220		12990			3270

$$\bar{X}=\frac{\sum_{i}^{n}x_i f_i}{\sum_{i}^{n}f_i}\ldots\ldots\ldots\ldots\ldots\ldots\bar{X}=\frac{x_1f_1+x_2f_2+\ldots\ldots\ldots\ldots+x_nf_n}{f_1+f_2+\ldots\ldots\ldots..+f_n}=\frac{12990}{220}=59.045\cong 59$$

$$M.D=\frac{\sum_{i=1}^{n}f_i\quad\left|x_i-\bar{X}\right|}{\sum_{i}^{n}f_{i}}=\frac{3270}{220}=14.864$$

واجب : البيانات التالية تمثل أوزان 50 طالب في إحدى الكليات

47	49	54	19	38	44	24	46	43	57	29	47
56	49	37	41	47	45	53	29	34	49	43	28
45	42	52	51	32	31	29	34	37	18	21	39
28	45	42	22	28	37	32	27	26	41	39	43
35	**23**										

كون جدول تكراري ثم اوجد قيمة كل من المدى والانحراف المتوسط لأوزان الطلبة.

3- الانحراف المعياري : يعرف بأنه الجذر ألتربيعي الموجب لمتوسط مجموع مربعات انحرافات قيم المتغير عن وسطه الحسابي ويرمز له بـ (S) .

أ- حساب الانحراف المعياري لبيانات غير مبوبة :

1- الطريقة المطولة (طريقة الانحرافات) :

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n}}$$

ملاحظة:يتم القســمة على n عندما تكون حجم العينةn اكبرمن 30. اي(n>30)وعندما تكون nاقل من 30.اي(n<30)فانه يقسم على (n-1)

مثال : البيانات التالية تمثل اوزان عينة من الطلبة عددها (10) المطلوب حســـاب قيمة الانحراف المعياري

56 , 68 , 78 , 63 , 65 , 68 , 71 , 69 , 62 , 56

$$\bar{X} = \frac{\sum_{i}^{n} x_i}{n} = \frac{650}{10} = 65$$

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n}}$$

$$S = \sqrt{\frac{(56-65)^2+(68-65)^2+(78-65)^2+(63-65)^2+(68-65)^2+(71-65)^2+(69-65)^2+(62-65)^2+(56-65)^2}{10}}$$

$$S = \sqrt{\frac{294}{10}} = \sqrt{29.4} = 5.422$$

مثال واجب : البيانات التالية تمثل أطوال عينة من الطلبة عددها (9) المطلوب حســـاب قيمة الانحراف المعياري.

120 , 150 , 185 , 191 , 172 , 175 , 177 , 165 , 185

ب- البيانات المبوبة :

1 – الطريقة المطولة :

$$S = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - \bar{X})^2}{\sum_{i}^{n} f_i}}$$

حيث ان f_i يمثل تكرارات الفئة

وان x_i مراكز الفئة

وان $\bar{X}$ يمثل الوسط الحسابي

مثال : الأتي توزيع تكراري لدرجات مجموعة من الطلبة بامتحان معين، احسب الانحراف المعياري لدرجات هؤلاء الطلبة باستخدام الطريقة المطلوبة

الفئات	f_i	x_i	$f_i x_i$	$(x_i - \bar{X})$	$(x_i - \bar{X})^2$	$f_i (x_i - \bar{X})^2$
0_10	4	5	20	-54	2916	11664
10_20	4	15	60	-44	1936	7744
20_30	8	25	200	-34	1156	9248
30_40	16	35	560	-24	576	9216
40_50	25	45	1125	-14	196	4900
50_60	58	55	3190	-4	16	928
60_70	42	65	2730	6	36	1512
70_80	35	75	2625	16	256	8960
80_90	18	85	1530	26	676	12168
90_100	10	95	950	36	1296	12960
المجموع	220		12990			79300

$$\bar{X} = \frac{\sum_{i}^{n} x_i f_i}{\sum_{i}^{n} f_i}$$

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \dots\dots\dots + x_n f_n}{f_1 + f_2 + \dots\dots\dots + f_n} = \frac{12990}{220} = 59.045 \cong 59$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - \bar{X})^2}{\sum_{i}^{n} f_i}} = \sqrt{\frac{79300}{220}} = \sqrt{360.454545} = 18.9856405$$

4- التباين : يعرف التباين بأنه متوسط مجموع مربعات انحرافات قيم المتغير عن وسطه الحسابي أي ان التباين هو مربع الانحراف المعياري ويرمز له (S^2)

أ- في حالة البيانات غير المبوبة :

1 – طريقة الانحرافات :

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n}$$

ب- في حالة البيانات المبوبة :

1 – طريقة الانحرافات :

$$S^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{X})^2}{\sum_{i}^{n} f_i}$$

واجب : اوجد قيمة التباين لجميع الامثله التي تم استخراج قيم الانحراف المعياري

خصائص الانحراف المعياري

1- ان $S \geq 0$ **و** $S^2 \geq 0$

2- اذا كانت (a) كمية ثابته وان $(y_i = ax_i)$ **فأن** $(s_y = |a| s_x)$ **و** $(s^2 y = |a| s^2 x)$

ولما كانت قيمة الانحراف المعياري لاي متغير قيمة موجبة، لذا يتم إهمال الإشارة السالبة، وعليه فان $s_y = as_x$.

مثال : اذا علمت ان $s_x = 6$ جد الانحراف المعياري الى $y = 4x$.

الحل :

$$s_y = |a| s_x = |4| .6 = 24$$

مثال : اذا علمت ان $s_x = 4$ جد الانحراف المعياري الى $y = -4x$.

الحل :

$$s_y = |a| s_x = |-3| .4 = 12$$

3- اذا كانت b كمية ثابتة حقيقية وان $y_i = x_i \pm b$ **عندئذ** $s_y = s_x$

مثال : اذا علمت ان $S_x = 6$ جد الانحراف المعياري الى $y = x + 3$

الحل:

$$S_y = S_x = 6$$

مثال : اذا علمت ان $S_x = 4$ جد الانحراف المعياري الى $y = x - 10$

الحل:

$$S_y = S_x = 4$$

4 – اذا كانت a, b ثوابت حقيقة وان $y_i = ax_i \pm b$ عندئذ

$$s_y = |a|.s_x$$

مثال : اذا علمت ان $S_x = 5$ جد الانحراف المعياري الى $y = 2x - 3$

الحل:

$$s_y = |a|.s_x$$

$$s_y = |2|.5 = 10$$

مميزات الانحراف المعياري

١- انه مقياس سهل الفهم والحساب.

٢- ان حسابه يستند الى كافة البيانات المتاحة.

٣- خضوعه للعمليات الجبرية.

٤- قابليته للتجزئة والاندماج .

عيوبه :

١- يتأثر وعلى نحو كبير باخطاء المعاينة.

٢- لا يمكن ايجاد قيمته في حالة التوزيعات التكرارية المفتوحة من طرف واحد او طرفين.

٣- تتأثر قيمته في حالة وجود قيم شاذة او متطرفة.

٤- لا يمكن ايجاد قيمته في حالة البيانات الوصفية.

ب- مقاييس التشتت النسبية :

وهي التي تبين درجة التجانس بشكل نسبي وتكون خالية من وحدات القياس ومن أهمها

١- **معامل الاختلاف :** قد يتطلب الامر في بعض الاحيان الى اجراء مقارنة بين تشتت مجموعتين او اكثر من القيم المختلفة عن بعضها من حيث الوسط الحسابي او ان قيم مفردات كل مجموعة مقاسة بوحدات قياس تختلف عن الاخرى ففي هذه الحالة فان مقياس التشتت المطلق سوف لن يكون مقياس نافع لوحده في اجراء مقارنات من هذا النوع انما يستوجب الامر ايجاد مقياس تشتت اخر اكثر ملائمة لهذه الحالات، هذا النوع من المقاييس يسمى بمقاييس التشتت النسبية واهم هذه المقاييس هو معامل الاختلاف ، ويرمز له (C.V) ويعبر عنه بالصيغة التالية :

$$C.V = \frac{S}{\overline{X}} \times 100$$

حيث ان :

S : تمثل الانحراف المعياري .

$\overline{X}$: الوسط الحسابي .

ملاحظة : كلما قلت قيمة معامل الاختلاف فان ذلك يدل على تجانس مفردات المتغير.

مثال : كان متوسـط درجات طلبة الصـف الاول في امتحان الرياضـيات (69) درجة وبانحراف معياري قدره (19.3) فيما كان متوسـط درجاتهم في امتحان الاحصــاء (75) وبانحراف معياري (25.5) في أي من الامتحانين كان مستوى اداء الطلبة اكثر تجانساً .

$$C.V = \frac{S}{\overline{X}} \times 100$$

$$C.V(math) = \frac{19.3}{69} \times 100 = 28\%$$

$$C.V = \frac{S}{\overline{X}} \times 100$$

$$C.V(stat) = \frac{25.5}{75} \times 100 = 34\%$$

حيث ان معامل الاختلاف في امتحان الرياضـيات اقل من معامل الاختلاف في امتحان الاحصاء ، عليه فان مستوى اداء الطلبة في امتحان الرياضيات كان اكثر تجانساً.

مثال : لدينا مجموعتين تم حساب الوسط الحسابي للمجموعة الأولى بالكيلو غرام والثانية بالباون وكانت نتيجة القياس ما يأتي:

المجموعة الأولى :	المتوسط الحسابي للوزن	70 كغم
	الانحراف المعياري	10 كغم
المجموعة الثانية :	المتوسط الحسابي للوزن	154.7 باون
	الإنحراف المعياري	22.1 باون

المطلوب : معرفة أي المجموعتين أكثر تجانساً (اقل تشتتاً):

الملاحظ هنا ان المجموعة الأولى أكثر تجانســـاً من المجموعة الثانية لأن الانحراف المعياري للمجموعة الأولى اصـــغر من الانحراف المعياري للمجموعة الثانية وهذا حكم خاطئ يحتاج إلى تحقيق من خلال معامل الاختلاف.

$$C.V = \frac{S}{\overline{X}} \times 100$$

$$C.V(1) = \frac{10}{70} \times 100 = 14.286\%$$

$$C.V = \frac{S}{\overline{X}} \times 100$$

$$C.V(2) = \frac{22.1}{154.7} \times 100 = 14.286\%$$

إذن المجموعتان متساويتان في التجانس أو التشتت لتساوي معامل الاختلاف على الرغم من اختلاف قيم الإنحراف المعياري وذلك بسـبب اختلاف وحدات القياس ويمكن مقارنة التشـتت من صفتين مختلفتين . مثلاً التشتت بين القوة العضلية والمطاولة العضلية

٢- الدرجة المعيارية:

في كثير من الاحيان نحتاج الى مقارنة قيمتين من قيم متغيرين مختلفين ففي هذه الحالة يجب تحويل وحدات هاتين القيمتين الى وحدات قياسية ويسمى بالدرجة القياسية (الدرجة المعيارية) ويرمز لها بالرمز (Z_i) وتأخذ الشكل التي :

$$Z_i = \frac{X_i - \overline{X}}{S}$$

حيث ان :

X_i : قيمة من المتغير X

S : تمثل الانحراف المعياري .

$\overline{X}$: الوسط الحسابي . $\times \sim (0,1)$

مثال : من المثال الاول الخاص في موضوع معامل الاختلاف اذا كانت درجة احد الطلبة في الرياضيات (76) ودرجته في الاحصاء (77) في أي مادة كان مستوى اداء هذا الطالب افضل .

الحل :

$$Z_i = \frac{X_i - \overline{X}}{S}$$

$$Z(MATH.) = \frac{76-69}{19.3} = 0.363$$

$$Z(STAT.) = \frac{77-75}{25.5} = 0.078$$

مستوى اداء الطالب في الرياضيات افضل من الاحصاء بالاعتماد النسبة الاكبر ، بينما كمقارنة مطلقة نلاحظ ان درجته في الإحصاء أعلى من درجته في الرياضيات.

الدرجة المعيارية المعدلة:

ملاحظة وجدت هذه الدرجة للتخلص من الحلات السالبة وهي درجة وسطها الحسابي (50) وانحرافها المعياري (10).

$$Z_o = \frac{X_i - \overline{X}}{S} * 10 + 50$$

لديك اختبار طلبة المرحلة الثانية بالسحب على العقلة والقفز من الثبات حصلوا على الدرجات التالية .

الاختبار	x	s
القفز على الثبات	190	4
السحب على العقلة	8	2

وقد حصلوا احد الطلاب على (180) سم بالقفزمن الثبات و7بالسحب على العقلة

المطلوب : في اي الاختبار يكون مستوى الطالب افضل

الفصل الرابع

مقياس الالتواء والتفلطح

لقد تناولنا في الفصول السابقة بعض المقاييس الوصفية التي تصف التوزيعات التكرارية وهما مقاييس التوسط ومقاييس التشتت.

في هذا الفصل سوف ندرس مقاييس اخرى تحدد شكل المنحني التكراري ومن حيث التماثل او الالتواء (Skewness) وتذبب القمة او تفطحلها (Kurtesis) .

١- مقاييس الالتواء

أ– معامل الالتواء (بطريقة المنوال).

ب– معامل الالتواء (بطريقة الوسيط).

ج– معامل الالتواء (بطريقة العزوم).

هذا وفي جميع هذه الطرق يكون الالتواء موجباً عندما يكون معامل الالتواء موجباً وسالباً عندما يكون معامل الالتواء سالباً ومتماثلاً عندما يكون معامل الالتواء صفراً.

أ– معامل الالتواء (بطريقة المنوال) ويرمز له بالرمز ($\propto_1$) ويكون خالي من الوحدات

$$\propto_1 = \frac{\bar{x} - M_\circ}{S}$$

$\bar{x}$: الوسط الحسابي

$M_\circ$: المنوال

S : الانحراف المعياري

فاذا كان $\bar{x} > M_\circ$ يكون $\bar{x} - M_\circ$ كمية موجبة نحصل على التواء الموجب (نحو اليمين).

فاذا كان $\bar{x} < M_\circ$ يكون $\bar{x} - M_\circ$ كمية سالب نحصل على التواء سالب (نحو اليسار).

اما اذا كان $\bar{x} = M_\circ$ نحصل على التواء او منحني متماثل .

ب– معامل الالتواء (بطريقة الوسيط) ويرمز له بالرمز ($\propto_2$) ويكون خالي من الوحدات

$$\propto_2 = \frac{3(\bar{x} - M_e)}{S}$$

$\bar{x}$: الوسط الحسابي

M_e : الوسيط

S : الانحراف المعياري

فاذا كان $M_e < \overline{x}$ نحصل على التواء السالب (نحو اليسار).

فاذا كان $M_e > \overline{x}$ نحصل على التواء موجب (نحو اليمين).

مثال : من البيانات التالية للجدول التكراري المبوبة

الفئات	f_i	x_i	$f_i x_i$	$(x_i - \overline{X})$	$(x_i - \overline{X})^2$	$f_i (x_i - \overline{X})^2$	x^2_i	$f_i x^2_i$
50-60	8	55	440	-24.7	610.09	4860.72	3025	24.200
60-	10	65	650	-14.7	216.09	2160.9	4225	42.250
70-	16	75	1200	-4.7	22.09	353.44	5625	90.000
80-	14	85	1190	5.3	28.09	393.26	7225	101.150
90-	10	95	950	15.3	234.09	2340.9	9025	90.250
100-	5	110	525	25.3	640.09	3200.45	11025	55.125
110-120	2	115	230	35.3	1246.09	2492.18	13225	26.450
	65		5185			15821085		429.425

احسب معامل الالتواء بطريقتي المنوال والوسيط

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{5185}{65} = 79.7$$

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

$$\frac{1582185}{64} = \sqrt{247.2} = 15.72$$

$$\bar{x} = 79.76$$

$$\mu_e = 79.06$$

$$\mu_0 = 7$$

$$\alpha_1 = \frac{\bar{x} - \mu_0}{s^2} = 0.14$$

$$\alpha_1 = \frac{\bar{x} - \mu_e}{s^2} = 0.14$$

اذاً $0 < 0.13$

المنحني ملتوي نحو اليمين (موجب)

الفصل الخامس
الارتباط

تناولنا سابقاً الوصف الإحصائي للمتغيرات من خلال مقاييس النزعة المركزية والتشتت وفي بعض الأحيان نحتاج الى وصف العلاقة بين متغيرين مثل القوة الانفجارية للرجلين والانجاز في الطفر العريض او حدت الانتباه ودقة التصويب في كرة السلة وهذا ما يدعى بالارتباط .

فالارتباط عبارة عن علاقة متبادلة بين متغيرين كميين او أكثر ، بحيث تؤدي زيادة او نقصـــان احدهما الى تغير مواز بالضـــرورة في المتغير الاخر، لذا فانه حينما يرتبط متغيران ارتباطاً عالياً ، فانه يكون من الممكن التنبؤ بقيم متغير معين، من خلال معرفة قيم المتغير الأخر.

أنواع الارتباط

١- الارتباط البسيط (بيرسون)

٢- الارتباط الرتبي (سبيرمان)..

معامل الارتباط

عبارة عن مؤشـــر عددي يســـتخدم للتعبير الكمي عن العلاقة الممتدة بين المتغيرين او اكثر، هذا المؤشر يسمى بمعامل الارتباط حيث يرمز لهذا المعامل في احصاء العينة بالرمز (r) وفي احصاء المجتمع الاصلي بالرمز (R).

انواع العلاقة؟

ا-العلاقة الطردية   أي كلمازاد(x) زاد(y) أو كلما قلت (x) قلت(y)

مثلا اذا زادت القوة العضـلية للذراعين ممكن ان يزاد الانجاز في رمي الثقل وهذه تسمى بالعلاقة الموجبة.

ب-العلاقة العكسية أي كلما زاد (x) قلت(y) أوكلما قلت (x) زاد(y)

مثلا اذا زاد وزن الجسم يمكن ان تقل اللياقة البدنية وهذه تسمى بالعلاقة العكسية.

ويتم التعبير عن هذا المعامل بالنسبة لإحصاء العينة كالتالي :

١- معامل الارتباط البســيط بيرســون (r_{xy}) : وهو يعني القيمة العددية لمقدار العلاقة بين المتغيرين (x ,y) ويســتخدم هذا المعامل لحســاب الارتباط البســيط بين مجموعتين من ازواج الدرجات لمتغيرين (x ,y).

طرق حساب معامل ارتباط بيرسون :

يتم حساب معامل ارتباط بيرسون باستخدام المعادلة التالية :

$$r_{xy} = \frac{\sum xy - n\bar{X}\bar{Y}}{\sqrt{(\sum_i^n x_i^2 - n\bar{X}^2)(\sum_i^n y_i^2 - n\bar{Y}^2)}}$$

ملاحظات :

١ - ان النتيجة التي تظهر لدينا تسمى بالقيمة المحتسبة وتتراوح القيمة ما بين $-1 \le r \le 1$

٢ - نقارن القيمة المحتسبة مع القيمة الجدولية فاذا كانت قيمة المحتسبة اكبر من الجدولية تسمى العلاقة معنوية واذا كانت اصغر من القيمة الجدولية تسمى العلاقة غير معنوية.

مثال : لدى اختبار (10) طلاب في السحب على العقلة حصلنا على النتائج آلاتية :

(2-7-3-5-4-6-7-8-3-1) ثم تم اختبارهم على جهاز المتوازي فحصلنا على النتائج الآتية (3-8-4-6-5-7-8-9-4-2) جد قيمة معامل الارتباط او ما هو العلاقة بين الاختبارين ؟

خطوات الحل :

١- نقوم بعمل جدول احصائي يتكون من 6 أعمدة و12 صفوف ، ثم نقوم بوضع أرقام الاختبارين في كل من x , y في الأعمدة وكما يأتي :

	x	y	xy	x^2	y^2
1	1	2	2	1	4
2	3	4	12	9	16
3	8	9	72	64	81
4	7	8	56	49	64
5	6	7	42	36	49
6	4	5	20	16	25
7	5	6	30	25	36
8	3	4	12	9	16
9	7	8	56	49	64
10	2	3	6	4	9
المجموع	$\sum x$	$\sum y$	$\sum xy$	$\sum x^2$	$\sum y^2$
	46	56	308	262	364

٢- نجد قيمة كل من الوسط الحسابي للاختبارين x , y وكما ياتي :

$$\bar{X} = \frac{\sum x}{n} = \frac{46}{10} = 4.6 \cong 5$$

$$\bar{Y} = \frac{\sum x}{n} = \frac{56}{10} = 5.9 \cong 6$$

$$r_{xy} = \frac{\sum xy - n\bar{X}\bar{Y}}{\sqrt{(\sum_i^n x_i^2 - n\bar{X}^2)(\sum_i^n y_i^2 - n\bar{Y}^2)}} = \frac{308-10(5)(6)}{\sqrt{(262-10(25))(364-10(36))}} = 1.154734 \cong 1$$

اذا الارتباط معنوي لان القيمة المحســوبة اكبر من الجدولية تحت مســتوى معنوي 0.05 =0.68عند درجات حرية (n-2)

ملاحظات :

١- الارتباط المعنوي : هو الارتباط حقيقي عند مستوى معنوي معين يكون المستوى 0.05 او 0.01 .

٢- الارتباط غير المعنوي : هو الارتباط العشوائي الذي يأتي نتيجة الصدفة.

٣- المستوى او النسبة الخطأ : بمعنى هي القيمة التي يكون فيها الارتباط معنوياً فاذا كانت نسبة الخطأ او المستوى المعنوي (0.05) بمعنى لو عدنا تجربة (100) نحصل على نتائج (95) مرة ونفشل 5 مرات في الحصول على الارتباط المعنوي.

مثال : حسب درجات اختبار الشد الى الأعلى واختبار ثني الذراعين من الجلوس المائل لـــــ4 طلاب فكانت النتائج كالتالي:

الشد الى الأعلى 6 , 5 , 7 , 9 واختبار ثني الذراعين 30 , 25 , 32 , 33

المطلوب حساب معامل ارتباط بيرسون بين درجات الاختبارين.علمابان القيمة الجدوالية 0.87 تحت مستوى معنوي عند درجات حرية (n-2)

خطوات الحل :

١- نقوم بعمل جدول إحصائي يتكون من 6 أعمدة و 5 صفوف، ثم نقوم بوضع أرقام الاختبارين في كل من x , y في الأعمدة وكما يأتي :

	X اختبار الشد الى الاعلى	Y اختبار ثني الذراعين	xy	x^2	y^2
1	9	33	297	81	1089
2	7	32	224	49	1024
3	5	25	125	25	625
4	6	30	180	36	900
المجموع	$\sum x$	$\sum y$	$\sum xy$	$\sum x^2$	$\sum y^2$
	27	120	826	191	3638

٢- نجد قيمة كل من الوسط الحسابي للاختبارين x , y وكما ياتي :

$$\bar{X} = \frac{\sum x}{n} = \frac{27}{4} = 6.75 \cong 7$$

$$\bar{Y} = \frac{\sum x}{n} = \frac{120}{4} = 30$$

$$r_{xy} = \frac{\sum xy - n\bar{X}\bar{Y}}{\sqrt{(\sum_i^n x_i^2 - n\bar{X}^2)(\sum_i^n y_i^2 - n\bar{Y}^2)}} = \frac{826-4(7)(30)}{\sqrt{191-4(46)(3638-4(900)}} = -14/\sqrt{190} = 1$$

واجب : قام احد الباحثين بحساب وزن الجسم بالأرطال ودرجات اختبار التعلق على العقلة من وضع مد الذراعين بالثواني لعدد ٥ طلاب في سن 12 سنة فكانت درجاتهم كالتالي

وزن الجسم x	80	117	96	85	92
زمن التعلق y	35	10	19	22	25

المطلوب حساب معامل الارتباط بين درجات وزن الجسم ودرجات اختبار التعلق على العقلة باستخدام معامل ارتباط بيرسون

٢- معامل الارتباط ألرتبي سبيرمان (r_s) : قد يتوفر لدينا بيانات على شكل وصفي او البيانات مختلطة (كمي ووصفي) وفي هذه الحالة نستخدم معامل الارتباط ألرتبي .

طرق حساب معامل الارتباط ألرتبي (سبيرمان) : يتم حســـــاب معامل الارتباط ألرتبي (سبيرمان) باستخدام المعادلة التالية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث ان d تمثل الفرق بين رتبت المتغير x عن رتبت المتغير y

مثال : اوجد معامل الارتباط ألرتبي (سبيرمان) لدرجات الطلبة عند التخرج من الإعدادية والتخرج من الجامعة والتي كانت على الشكل آلاتي :

x	y
(9)56	مقبول(8)
(8)60	متوسط(6)
(5)70	متوسط عالي(5)
(2)85	جيد(3)
(4)80	جيد جدا(2)
(1)90	امتياز(1)
(3)82	جيد(4)
(6)70	متوسط(7)
(7)66	مقبول(8)

الحل :١- لتسهيل عملية الحل نرتب البيانات تصاعدياً او تنازلياً

٢ لحصل على رتب كل متغير وكما يأتي :

رتب x	رتب y	d	d^2
9	8.5	0.5	0.25
8	6.5	1.5	2.25
5.5	5	0.5	0.25
2	3.5	-1.5	2.25
4	2	2	4
1	1	0	0
3	3.5	-0.5	0.25
5.5	6.5	-1	1
7	8.5	-1.5	2.25
المجموع			12.5

٣- نطبق القانون ونجد قيمة r

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6*12.5}{9(81-1)} = 0.896$$

هذا يعني ان هناك علاقة بين درجات الإعدادية ودرجات التخرج في الجامعة للطالب

مثال : في الجدول الأتي تقديرات ستة لاعبين في إحدى فعاليات الجمباز إذ قام بوضع التقديرات حكمان ، والمطلوب حساب معامل الارتباط (سبيرمان) بين رأي الحكمان.

رقم اللاعب	١	٢	٣	٤	٥	٦
رأي الحكم الأول (x)	ضعيف(**5**)	ممتاز(**1**)	جيد(**3**)	ضعيف جداً(**6**)	مقبول(**4**)	جيد جداً(**2**)
رأي الحكم الثاني (y)	مقبول(**4**)	جيد جداً(**2**)	جيد(**3**)	ضعيف(**5**)	ضعيف جداً(**6**)	ممتاز(**1**)

الحل : ١- لتسهيل عملية الحل نرتب البيانات تصاعدياً او تنازلياً

٢-لحصل على رتب كل متغير وكما يأتي

رتب x	رتب y	d	d^2
5	4	1	1
1	2	-1	1
3	3	0	0
6	5	1	1
4	6	-2	4
2	1	1	1
المجموع			8

٣- نطبق القانون ونجد قيمة r

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6*8}{6(36-1)} = 0.77$$

هذا يعني ان هناك علاقة بين تقدير الحكمان

واجب : جرت بطولة القطر في بناء الأجسام ، وشارك فيها عشرة لاعبين، قام بالتحكيم حكمان، قيمة الدرجة الكبرى (20) وكانت الدرجات التي وضعها الحكمان كما في الجدول الأتي :

درجة الحكم y	درجة الحكم x	رقم اللاعب
ضعيف جداً	15	1
جيد	16	2
ضعيف	14	3
ضعيف جداً	11	4
جيد جداً	20	5
ضعيف جداً	13	6
ضعيف جداً	12	7
ممتاز	19	8
مقبول	17	9
جيد	18	10
		مجموع

الفصل السادس

اختبار دلالة الفروق بين المتوسطات

اختبار (ت- t) لدلالة الفروق بين المتوسطات

يعد اختبار T. Test من أكثر اختبارات الدلالة شيوعاً في الأبحاث النفسية والتربوية والرياضية. وترجع نشأته إلى أبحاث (ستودنت) ولهذا سمي بأكثر الحروف تكراراً في اسمه وهو الحرف (ت)، ويستخدم اختبار (t) لقياس دلالة فروق المتوسطات غير المرتبطة والمرتبطة، للعينات المتساوية وغير المتساوية، والعينات الصغيرة والعينات الكبيرة وعند استخدام اختبار (t) لدلالة الفروق بين المتوسطات على الباحث أن يدرس خصائص متغيرات البحث من النواحي الآتية:

١- حجم العينة.

٢- الفرق بين عينتي البحث.

٣- مدى تجانس العينة.

وتتلخص الحالات المختلفة لحساب دلالة الفروق بين المتوسطات باختبار (t) فيما يأتي:

أولاً - دلالة الفروق لمتوسطين غير مرتبطين ولعينتين متساويتين ويمكن إيجاده على وفق المعادلة الآتية:

نستخدم القانون الاتي :

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2 + S_2^2}{n-1}}}$$

مثال : عينتان مستقلتان متساويتان في العدد ، العينة الأولى (n=7) سحبت من مجتمع من الرياضـــيين، والعينة الأخرى (n=7) ســـحبت من مجتمع من غير الرياضـــيين، وعندما حسبت درجات نسبة الذكاء باستخدام اختبار ستانفورد ، كانت البيانات بالنسبة للعينتين كالتالي:

مجموعة الرياضيين	مجموعة غير الرياضيين
n = 7	n = 7
$\bar{X}_1 = 110.4$	$\bar{X}_2 = 97.9$
$S_1 = 10.81$	$S_2 = 8.46$

المطلوب : اختبار دلالة الفرق بين متوســـطى العينتين (اختبار t) علماً بان القيمة الجدولية تساوي 3.06

$$t=\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{S_1^2+S_2^2}{n-1}}}=\frac{110.4-97.9}{\sqrt{\frac{(10.81)^2+(97.9)^2}{7-1}}}=2.23$$

مثال واجب : قام باحث بأجراء دراسة مقارنة المستوى القوة العضلية بين شعبتين (أ،ب) طلاب المرحلة الرابعة، اجري اختبار السـحب على العقلة وحصـلوا على التكرارات التالية علماً بان القيمة الجدولية تساوي 1.83 .

أ:	**7**	**7**	**4**	**5**	**9**	**8**	**4**	**5**	**7**	**6**
ب :	**8**	**7**	**9**	**8**	**9**	**8**	**4**	**6**	**7**	**5**

ثانياً - دلالة الفروق لمتوسطين غير مرتبطين ولعينتين غير متساويتين ويمكن إيجاده على وفق المعادلة الآتية:

$$t=\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\left[\frac{nS_1^2+nS_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$$

مثال : عينتان عشـوائيتان تم اختيار كل منهما من مجتمع يتوزع توزيعاً طبيعياً ببياناتهم وكالتالي :

العينة الأولى ($n_1=15$) طالباً تم اختيارهم عشوائياً من أحدى كليات الطب.

العينة الثانية ($n_2=18$) طالباً تم اختيارهم عشوائياً من أحدى كليات الفنون.

وقد طبق اختبار للذكاء على العينتين ، فكانت درجاتهم كالتالي :

العينة الاولى	العينة الثانية
n = 15	n = 18
$\overline{X}_1=120$	$\overline{X}_2=110$
$S_1^2=100$	$S_2^2=64$

المطلوب : اختبار الفرض الذي يقرر ان طلاب كلية الطب لهم معامل ذكاء يفوق معامل ذكاء طلاب كلية الفنون علماً بان قيمة (t) الجدولية تساوي 2.042.

الحل : بما ان العينتين غير مرتبطتين (مستقلتين) وغير متساويتين بالحجم نتبع القانون التالي :

$$t=\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\left[\frac{nS_1^2+nS_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$$

$$t=\frac{120-110}{\sqrt{\left[\frac{15(100)+18(64)}{15+18-2}\right]\left[\frac{1}{15}+\frac{1}{18}\right]}}=3.189$$

بما ان قيمة (t) المحســـوبة اكبر من قيمة (t) الجدولية هذا يعني انه يوجد فرق بين متوســـطى ذكاء المجتمعين الإحصـائيين اللذين سـحبت منهما العينتان أي انه متوسـط ذكاء المجتمع الأول يختلف عن متوسط ذكاء المجتمع الثاني.

مثال : البيانات التالية لاختبار بدني طبق على مجموعتين مســـتقلتين من الطلاب، أحداهما من الممارسين

للنشاط الرياضي والأخرى من غير الممارسين للنشاط الرياضي وكالتالي :

مجموعة الممارسين	مجموعة الغير ممارسين
n = 17	n = 15
$\overline{X}_1 = 80$	$\overline{X}_2 = 75$
$S_1^2 = 16$	$S_2^2 = 18$

المطلوب : اختبار دلالة الفرق بين متوسطي المجموعتين (اختبار t).

الحل : بما ان العينتين غير مرتبطتين (مستقلتين) وغير متساويتين بالحجم نتبع القانون التالي :

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left[\frac{nS_1^2 + nS_2^2}{n_1 + n_2 - 2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

$$t = \frac{80 - 75}{\sqrt{\left[\frac{17(16) + 18(18)}{16 + 18 - 2}\right]\left[\frac{1}{16} + \frac{1}{18}\right]}} = 3.466$$

بما ان قيمة (t) المحســـوبة اكبر من قيمة (t) الجدولية هذا يعني انه يوجد فرق بين متوســـطى المجتمعين الإحصائيين اللذين سحبت منهما العينتان

واجب : طبق اختبار الثقة بالنفس على المجموعتين احدهما من اللاعبين والبالغ عددهم 40 لاعباً والمجموعة الأخرى من اللاعبات وعددهم 28 لاعبة وكانت نتائج الاختبار على الشـــكل التالي علماً بان القيمة الجدولية تساوي 1.64.

عينة اللاعبين	عينة اللاعبات
n = 40	n = 28
$\overline{X}_1 = 36.82$	$\overline{X}_2 = 33.71$
$S_1 = 4.63$	$S_2 = 4.94$

المطلوب : اختبار دلالة الفرق بين متوسطي المجموعتين (اختبار t).

ثالثاً - دلالة الفروق لمتوسطين مرتبطين (غير مستقلة) ولعينتين متساويتين ويمكن إيجاده على وفق المعادلة الآتية:

$$tv = \frac{\sum vi}{\sqrt{n(\sum v^2 i) - (\sum vi)^2 \quad / n-1}}$$

حيث ان vi يمثل الفرق بين متوسط الاختبار القبلي والاختبار البعدي

وان viمجموع مربعات الفروق بين درجاتالقياس القبلي و البعدي

ملاحظة:يجب ان تكون n<30 فيكون توزيع معتدل.فاذا كانت tv بين القياســين البعدي والقبلي اذاكان موجبة دل على ان الفرق لصــالح القياســي القبلي .اذا كانت ســالبة دل ذلك ان الفروق لصالح القياسي البعدي

مثال : اجري اختبار الوثب الطويل من الثبات لعدد من الطلاب وكانت نتيجة متوســط الاختبار القبلي والمبين في الجدول أدناه وبعد أجراء البرنامج التدريبي لمدة ثلاثة أشـــهر اجري الاختبار ألبعدي لنفس الطلاب وكانت نتيجة متوســـط الاختبار وكما مبينه بالجدول أدناه وان قيمة (t) الجدولية تساوي 1.83

الطلاب	الاختبار القبلي	الاختبار البعدي	vi	$\vec{v}^2$
1	160	180	-20	400
2	150	190	-40	1600
3	170	200	-30	900
4	180	210	-30	900
5	160	200	-40	1600
6	150	190	-40	1600
7	170	180	-10	100
8	180	190	-10	100
9	160	180	-20	400
10	180	200	-20	400
المجموع			-260	8000

$$tv = \frac{\sum v}{\sqrt{n\sum v^2 - (\sum v)^2 \quad / n-1}}$$

$$tv = \frac{-260}{\sqrt{10(8000) - (260)^2 \quad / 9}} = \frac{-260}{\sqrt{80000 - 67600 / 9}} = \frac{-269}{37.1184} = -7.005$$

بمان tv المحسوبة اكبر tv الجدولية اذا يوجد فروق معنوية لصالح البعدي.

واجب : قام باحث بوضع برنامج تدريبي للارتقاء بصفة مطاولة القوة العضلية للذراعين لعينة قدرها (١٥) لاعباً ، إذ تم اختبارهم وحصلوا على التكرارات الآتية(٣، ٤، ٣، ٥، ٤، ٦، ٧، ٦، ٥، ٤، ٨، ٣، ٢، ٤، ٥) ثم طبق البرنامج التدريبي الذي استغرق شهرين ثم أعيد الاختبار ألبعدي على العينة نفسها وحصلوا على الدرجات الآتية : (٧، ٦، ٨، ٦، ٩، ١٠، ٧، ٨، ٦، ٨، ٩، ١٠، ٩، ٨، ١١) .

المطلوب: هل هناك فروق بين الاختبارين القبلي والبعدي للعينة في اختبار القوة العضلية للذراعين؟

٢- التحليل التباين (ANOVA)

لنفرض ان لدينا ثلاث مجموعات من الأفراد، ثم قمنا بقياس انجازهم في العمل في ظل ظروف مختلفة المستوى وفقاً لمتغير الضغوط النفسية، حيث تمثل المستويات المختلفة للضغوط النفسية المتغير المستقل او متغير المعالجات، فإذا قسنا درجات نجاح هؤلاء الأفراد في العمل (كمتغير تابع) ، وكانت درجاتهم كالتالي :

مستويات الضغوط النفسية	متوسطات الاداء
منخفض	$\overline{X}_1 = 10$
متوسط	$\overline{X}_2 = 20$
مرتفع	$\overline{X}_3 = 30$

فانه لا يجوز في مثل هذه الحالة دراسة تأثير الضغوط النفسية على مستويات الانجاز عن طريق تحليل الفروق بين مجموعتين فقط وذلك لكون المتغير المستقل متعدد المستويات، وانه يتوقع ان يكون المتغير التابع ايضاً متعدد المستويات.

في هذه الحالة نستخدم تحليل التباين (ANOVA) لعامل واحد لتقويم تأثير متغير مستقل مفرد له أكثر من مستويين على متغير، ومن امثلة المتغيرات المستقلة متعددة المستويات في مجال التربية البدنية والرياضية ما يلي :

اللياقة البدنية : مرتفع متوسط منخفض

المهارة في كرة القدم : مرتفع متوسط منخفض

التدريب : لايتدرب مرة في الاسبوع ٣ مرات في الاسبوع ٥ مرات

الشدة في التدريب : خفيف متوسط اقل من القصوى

الرموز المستخدمة في تحليل التباين ودلالاتها :

يستخدم في تحليل التباين لعامل واحد الرموز التالية :

$$1 - C = \frac{(\sum_{1}^{n} y_i)^2}{\sum_{1}^{n} n_i}$$ معامل التصحيح

$$2 - y_{..} = \sum_{1}^{n} {y_i}^2 - C = SST$$ المجموع الكلي

$$3 - SST = SS_t + SS_r$$

SST : المجموع الكلي للمربعات وهو يمثل مربعات انحرافات كل المشاهدات (القيم) عن المتوسط الحسابي العام.

SS_t : مجموع المربعات بين المجموعات ويعرف باسم مجموع مربعات المعالجات .

$$SS_t = \frac{(\sum y_1)^2}{n_1} + \frac{(\sum y_2)^2}{n_2} + \frac{(\sum y_3)^2}{n_3} + \frac{(\sum y_4)^2}{n_4} - C$$

SS_r : مجموع المربعات داخل المجموعات ويعرف باسم مجموع مربعات الخطأ .

$$SS_r = SST - SS_t$$

ويستخدم لحساب تحليل التباين في اتجاه واحد المعادلات التالية :

$$4 - MS_t = \frac{SS_t}{dF_t}$$

حيث ان :

MSt : متوسط مربع المعالجات (بين المجموعات).

SSt : مجموع المربعات بين المجموعات.

dFt : درجات الحرية (عدد المجموعات -١).

$$5 - MS_r = \frac{SS_r}{dF_r}$$

حيث ان :

MSr : متوسط مربع الخطأ (داخل المجموعات).

SSr : مجموع المربعات داخل المجموعات.

dFr : درجات الحرية (عدد الأفراد) - (عدد المجموعات-١).

$$6 - F = \frac{MS_t}{MS_r}$$

حيث ان :

F : نسبة (ف-F) المحسوبة.

مثال : أراد احد الباحثين أجراء تجربة للتعرف على تأثير التدريب العقلي والتدريب البدني والتدريب العقلي والبدني معاً على تعلم مهارة حركية ما، فقام باختبار أربع مجموعات تتكون كل مجموعة من (6) أفراد، ثم قام بتوزيع المجموعات عشوائياً على برامج التدريب التي استمرت لمدة أربعة أسابيع بواقع (30) دقيقة يومياً لكل مجموعة من المجموعات التجريبية الثلاثة ، في حين لم تتدرب المجموعة الضابطة على آية برامج، وبعد انتهاء المدة المقررة للتدريب على البرامج الثلاثة قام الباحث باختبار مجموعات البحث الأربع في المهارة الحركية، وقد توصل الى النتائج التالية :

المجموعة الاولى	المجموعة الثانية	المجموعة الثالثة	المجموعة الرابعة
تدريب عقلي	تدريب عقلي وبدني	تدريب بدني	ضابطة
16.0	17.5	17.5	15.0
17.17	19.0	18.5	16.5
17.5	19.5	19.5	17.0
18.5	21.0	20.5	18.0
19.0	21.0	21.5	18.0
17.5	20.0	19.5	16.0
N=6	N=6	N=6	N=6
$\overline{X}=17.61167$	$\overline{X}=19.66667$	$\overline{X}=19.5$	$\overline{X}=16.75$

<u>المطلوب :</u> حساب قيمة F ، ثم اختبار الدلالة الإحصائية لقيمة F المحسوبة.علماً بان قيمة F الجدولية تساوي 3.10 ثم تأكد من معنوية الفروق ارسم جدول إحصائي.

الحل : ١- نقوم بصياغة الفرض الصفري والفرض البديل كالتالي :

$H_0 \qquad \overline{X}_1 = \overline{X}_2 = \overline{X}_3$

$H_1 \qquad \overline{X}_1 \neq \overline{X}_2 \neq \overline{X}_3$

٢- تحديد مستوى المعنوية 0.05 حيث نلاحظ الأتي:

ترفض الفرضية الصفرية H_0 اذا كانت ف-F المحسوبة $\leq$ F الجدولية.

٢- نجد معامل التصحيح

$$C = \frac{(\sum_{1}^{n} y_i)^2}{\sum_{1}^{n} n_i}$$

	Y_1	Y_2	Y_3	Y_4	Y_1^2	Y_2^2	Y_3^2	Y_4^2	$\bar{Y}$
1	**16.0**	**17.5**	**17.5**	**15**	**256**	**306.25**	**306.25**	**225**	**16.5**
2	**17.17**	**19.0**	**18.5**	**16.5**	**294.8089**	**361**	**342.25**	**272.25**	**17.7925**
3	**17.5**	**19.5**	**19.5**	**17**	**306.25**	**380.25**	**380.25**	**289**	**18.375**
4	**18.5**	**21.0**	**20.5**	**18**	**342.25**	**441**	**420.25**	**324**	**19.5**
5	**19.0**	**21.0**	**21.5**	**18**	**361**	**441**	**462.25**	**324**	**19.875**
6	**17.5**	**20.0**	**19.5**	**16**	**306.25**	**400**	**380.25**	**256**	**73**
المجموع	**105.67**	**118**	**117**	**100.5**	**1866.559**	**2329.5**	**2291.5**	**1690.25**	**165.0425**

$$C = \frac{(\sum_{1}^{n} y_i)^2}{\sum_{1}^{n} n_i} = \frac{(105.67+118+117+100.5)^2}{24} = 8109.624$$

$$2 - SST = y_{..} = \sum_{1}^{n} y_i^2 - C = 1866.559 + 2329.5 + 2291.5 + 1690.25 - 8109.624 = 8177.809 - 8109.624 = 68.1852$$

$$SS_t = \frac{(\sum y_1)^2}{n_1} + \frac{(\sum y_2)^2}{n_2} + \frac{(\sum y_3)^2}{n_3} + \frac{(\sum y_4)^2}{n_4} - C$$

$$= \frac{(105.67)^2}{6} + \frac{(118)^2}{6} + \frac{(117)^2}{6} + \frac{(100.5)^2}{6} - 8109.624$$

$$= 8146.566 - 8109.624 = 36.94278$$

$$3 - SST = SS_t + SS_r \Rightarrow SS_r = SST - SS_t$$

$$SS_r = 68.1852 - 36.94278 = 31.24242$$

dFt : درجات الحرية (عدد المجموعات - ١).

$$dFt = 4 - 1 = 3$$

$$4 - MS_t = \frac{SS_t}{dF_t} = \frac{36.94278}{3} = 12.31426$$

dFr : درجات الحرية (عدد الافراد) - (عدد المجموعات-١).

$$dFr = (24) - (4 - 1) = 21$$

$$5 - MS_r = \frac{SS_r}{dF_r} = \frac{31.24242}{21} = 1.48773$$

$$6-F=\frac{MS_t}{MS_r}=\frac{12.31426}{1.48773}=8.27721$$

نسبة F	التباين MS	درجة الحرية dF	مجموع المربعات SS	مصدر التباين
8.27721	**12.31426**	**3**	SS_t **36.94278**	بين المجموعات (المعالجات)
	1.48773	**21**	SS_r **31.24242**	داخل المجموعات (الخطا)
	——	**24**	SST 68.1852	التباين الكلي

القرار الإحصائي:

بما ان F المحسوبة < F الجدولية

7.883039 < 3.10 اذا يرفض الفرضية الصفرية الذي يقرر ان متوسطات المجموعات الثلاث متساوية، مما يعني ان برامج التدريب الثلاثة احدثث فروقا دالة احصائيا بين درجات المجموعات الاربع.

<u>مثال :</u> اجري اختبار السحب على العقلة لثلاثة مجاميع (A,B,C) بحيث ان كل مجموعة تتكون من (5) لاعبين وحصـــــلوا على التكرارات التالية ، المطلوب أيجاد الفروق بين المجاميع الثلاثة علماً بان قيمة F الجدولية تساوي 3.88 تم تأكد من معنوية الفروق مع رسم جدول إحصائي.

A	B	C
5	7	3
2	7	4
6	7	7
5	8	3
4	9	2
N=5	N=5	N=5
$\overline{X}=22$	$\overline{X}=38$	$\overline{X}=19$

<u>المطلوب :</u> حساب قيمة F ، ثم اختبار الدلالة الإحصـائية لقيمة F المحسوبة.علماً بان قيمة F الجدولية تساوي 3.10 ثم تأكد من معنوية الفروق ارسم جدول إحصائي.

الحل : ١- نقوم بصياغة الفرض الصفري والفرض البديل كالتالي :

$$H_0 \qquad \overline{X}_1 = \overline{X}_2 = \overline{X}_3$$
$$H_1 \qquad \overline{X}_1 \neq \overline{X}_2 \neq \overline{X}_3$$

٢- تحديد مستوى المعنوية 0.05 حيث نلاحظ الأتي:

ترفض الفرضية الصفرية H_0 اذا كانت ف-F المحسوبة ≤ F الجدولية.

٣ - نجد معامل التصحيح

$$1-C = \frac{(\sum_{1}^{n} y_i)^2}{\sum_{1}^{n} n_i}$$

	Y_1	Y_2	Y_3	Y_1^2	Y_2^2	Y_3^2	$\bar{Y}$
1	**3**	**5**	**7**	**9**	**25**	**49**	**5**
2	**4**	**2**	**7**	**16**	**4**	**49**	**4.333333**
3	**7**	**6**	**7**	**49**	**36**	**49**	**6.666667**
4	**3**	**5**	**8**	**9**	**25**	**64**	**5.333333**
5	**2**	**4**	**9**	**4**	**16**	**81**	**5**
المجموع	**19**	**22**	**38**	**87**	**106**	**292**	**26.33333**

$$C = \frac{(\sum_{1}^{n} y_i)^2}{\sum_{1}^{n} n_i} = \frac{(19+22+38)^2}{15} = 416.06$$

$$2 - SST = y_{..} = \sum_{1}^{n} {y_i}^2 - C = 106 + 292 + 87 - 416.06 = 68.94$$

$$SS_t = \frac{(\sum y_1)^2}{n_1} + \frac{(\sum y_2)^2}{n_2} + \frac{(\sum y_3)^2}{n_3} - C$$

$$= \frac{(22)^2}{5} + \frac{(38)^2}{5} + \frac{(19)^2}{5} - 416.06$$

$$= 457.8 - 416.06 = 41.74$$

$$3 - SST = SS_t + SS_r \Rightarrow SS_r = SST - SS_t$$

$$SS_r = 68.94 - 41.74 = 27.2$$

dFt : درجات الحرية (عدد المجموعات – ١).

$$dFt = 3 - 1 = 2$$

$$4 - MS_t = \frac{SS_t}{dF_t} = \frac{41.74}{2} = 20.87$$

dFr : درجات الحرية (عددٍ الإفراد – ١) – (عدد المجموعات–١).

$$dFr = (15) - (3 - 1) = 13$$

$$5 - MS_r = \frac{SS_r}{dF_r} = \frac{27.2}{13} = 2.09$$

$$6 - F = \frac{MS_t}{MS_r} = \frac{20.87}{2.09} = 9.99$$

مصدر التباين	مجموع المربعات SS	درجة الحرية dF	التباين MS	نسبة F
بين المجموعات (المعالجات)	SS_t **41.74**	**2**	**20.86**	**9.99**
داخل المجموعات (الخطا)	SS_r **27.2**	**13**	**2.09**	
التباين الكلي	SST 68.94	**15**	——	

القرار الإحصائي:

بما ان F المحسوبة < F الجدولية

9.23 < 3.10 اذا يرفض الفرضـــية الصـــفرية الذي يقرر ان متوســـطات المجموعات الثلاث متســاوية، مما يعني ان الاختبار على المجموعة الثلاثة احدثث فروقا دالة احصــائيا بين درجات المجموعات الثلاثة.

واجب : اراد احد الباحثين اجراء تجربة لمقارنة التاثير النســـبي لبرامج التدريب الايزومتري والايزوتوني للقوة العضلية، فقام باختبار مجموعة تتكون من 20 طالباً من طلاب المرحلة الثانوية للذكور، ثم قام بتوزيعهم عشوائياً على ثلاث مجموعات ، المجموعة الاولى N=7 والمجموعة الثانية N=8 والمجموعة الثالثة N=5 ، ثم قام باعداد برنامج للتدريب الايزومتري للمجموعة الاولى، واعداد برنامج اخر للتدريب الايزوتوني للمجموعة الثانية، وابقى المجموعة كمجموعة ضابطة، وبعد التدريب لمدة 8 اسابيع على برامج التدريب المقترحة قام بتطبيق اختبار الشد لاعلى على المجموعات الثلاث، وقد حصل على البيانات التالية :

التدريب الايزومتري N=7	التدريب الايزوتونى N=8	الضابطة N=5
10	15	4
8	10	7
12	15	6
14	11	8
13	11	6
11	9	
16	13	
	11	

المطلوب : حساب قيمة F ، ثم اختبار الدلالة الإحصائية لقيمة F المحسوبة.علماً بان قيمة F الجدولية تساوي 17.2 ثم تأكد من معنوية الفروق ارسم جدول إحصائي.

٣- المقارنات المتعددة

اشرنا فيما سبق ان تحليل التباين لعامل واحد يستخدم في حالة العينات التي تزيد عن اثنين وذلك بغرض التعرف على ما اذا كان المتغير المستقل له تاثير على المتغير التابع.

ففي الامثلة الســـابقة يظهر ان قيمة F المحســـوبة ذات دلالة معنوية مما يدل على ان المعالجات الخاصة بالمتغير المستقل وهو التدريب بمستوياته المختلفة كان له تأثير جوهري على المجاميع في حين ان قيمة F المحســـوبة من تحليل التباين لم تخبرنا أي برامج التدريب كان اكثر فعالية عن غيرها ، مما يســـتلزم في مثل هذه الحالة أجراء المقارنات بين المتوســـطات الحسابية التي حصلنا عليها من اختبار F لاستخلاص معلومات اكثر عن متغيرات الدراسة عن طريق المقارنات المتعددة والتي تستخدم في التجارب البحثية بأسلوبين رئيسيان هما :

- **المقارنات المتعددة القبلية**
- **المقارنات المتعددة البعدية.**

ويســـتخدم في المقارنات المتعددة في مجال علوم التربية البدنية والرياضـــية عدد من الطرق من اهمها ما يلي :

١- **طريقة اقل فرق معنوي (L.s.d).**
٢- **اختبار (t) لدونتي.**
٣- **اختبار دانكان**

وغيرها من الطرق وسوف نقوم بدراسة الطريقة الأولى لضيق الوقت فقط.

طريقة اقل فرق معنوي (L.s.d).

تعتبر هذه الطريقة من اكثر الطرق شـــائعة الاســـتخدام بغرض تحديد الفروق الحقيقة بين متوسطات المعالجات في التجارب متعددة المجموعات لمعرفة اقل فرق معنوي (L.s.d).

ويهدف حســـاب اقل فرق معنوي (L.s.d) تقديم اقل قيمة يمكن قبولها لكي يكون الفرق بين متوسطى عينتين (مجموعتين) دالا إحصائيا ، ويلاحظ ان من شروط هذه الطريقة الا تستخدم اذا كانت F المحســـوبة غير دالة احصـــائياً ، أي ان اســـتخدام هذه الطريقة مشـــروط بالدلالة الإحصائية لنسبة F المحسوبة من التجربة.

المعادلة الأساسية لهذه الطريقة :

نســتخدم لحســاب اقل فرق معنوي (L.s.d) بالنســبة للمجموعات المتســاوية العدد المعادلة التالية :

$$L.s.d = t_{(N,0.05 or 0.01)} \times \sqrt{\frac{2 \times Msr}{N}}$$

حيث ان :

MSr : تباين الخطأ.

$t_{(N,0.05 or 0.01)}$: قيمة (t) الجدولية عن مستوى دلالة (0.05 او 0.01).

N : عدد الملاحظات (المفردات) في كل عينة من العينات.

مثال : في مثال تجربة التعرف على تأثير التدريب العقلي والتدريب البدني والتدريب العقلي والبدني معاً اظهرت نتائج التحليل التباين ان F المحسوبة > F الجدولية عند مستوى (0.05) 7.883039 > 3.10 ، حينئذ نقوم بمقارنة المتوســط الحســابي لكل مجموعة تجريبية على حدة مع المجموعة الضــابطة عند مستوى (0.05) ايضـاً ، وذلك بهدف التعرف على التأثير النسـبي لنوع التدريب على تعلم المهارات الحركي.

بما ان قيمة كل من

$$MS_r = 1.562121$$

المجوعات متساوية العدد وتساوي (6).

$t_{(20,0.05)} = 2.09$: قيمة (t) الجدولية عن مستوى دلالة (0.05) وذلك بالرجوع الى جدول توزيع (t) عند مستوى (0.05) ودرجة حرية تساوي درجة حرية الخطأ والتي تساوي (20).

نقوم بالتعويض بمعادلة **اقل فرق معنوي (L.s.d) بالنسبة للمجموعات المتساوية العدد المعادلة التالية :**

$$L.s.d = t_{(20,0.05)} \times \sqrt{\frac{2 \times Msr}{N}}$$

$$= 2.09 \times \sqrt{\frac{2 \times 1.54}{6}}$$

$$= 2.09 \times \sqrt{\frac{3.08}{6}} = 1.50$$

بعد ذلك بين متوسطات الحسابية لكل مجموعة من المجموعة الضابطة.

	Y_1	Y_2	Y_3	Y_4	Y_1^2	Y_2^2	Y_3^2	Y_4^2	$\bar{Y}$
1	16.0	17.5	17.5	15	256	306.25	306.25	225	16.5
2	17.17	19.0	18.5	16.5	294.8089	361	342.25	272.25	17.7925
3	17.5	19.5	19.5	17	306.25	380.25	380.25	289	18.375
4	18.5	21.0	20.5	18	342.25	441	420.25	324	19.5
5	19.0	21.0	21.5	18	361	441	462.25	324	19.875
6	17.5	20.0	19.5	16	306.25	400	380.25	256	73
المجموع	105.67	118	117	100.5	1866.559	2329.5	2291.5	1690.25	165.0425
الوسط الحسابي	17.61167	19.66667	19.5	16.75	311.0932	388.25	381.9167	281.7083	27.50708

$\bar{X}_1 - \bar{X}_4 = 17.61167 - 16.75 = 0.861667$

$\bar{X}_2 - \bar{X}_4 = 19.6667 - 16.75 = 2.916667$

$\bar{X}_3 - \bar{X}_4 = 19.5 - 16.75 = 2.75$

القرار :

كلاً من المجموعة الثانية والثالثة اكبر من قيمة (L.s.d) أي ان $0.861667 \text{and } 2.916667 > 1.50$ **وان المجموعة الأولى اصـغر من قيمة (L.s.d) أي ان** $2.75 < 1.50$ **هذا يدل على ان التدريب البدني عند ربطه مع التدريب العقلي يؤدي الى تحسين التعلم الحركي، في حين ان التدريب العقلي بمفرده لم يكن له تأثير يذكر.**

<u>مثال :</u> اختبار السـحب على العقلة لثلاثة مجاميع (A,B,C) بحيث ان كل مجموعة تتكون من (5) لاعبين وحصلوا على التكرارات التالية ، <u>المطلوب أي المجموعة أفضل</u> .

A	B	C
5	7	3
2	7	4
6	7	7
5	8	3
4	9	2
N=5	N=5	N=5
$\bar{X} = 22$	$\bar{X} = 38$	$\bar{X} = 19$

بما ان قيمة كل من

$$MS_r = 2.26$$

المجوعات متساوية العدد وتساوي (5).

$t_{(12,0.05)} = 1.78$: قيمة (t) الجدولية عن مستوى دلالة (0.05) وذلك بالرجوع إلى جدول توزيع (t) عند مستوى (0.05) ودرجة حرية تساوي درجة حرية الخطأ والتي تساوي (12).

نقوم بالتعويض بمعادلة **اقل فرق معنوي (L.s.d) بالنسبة للمجموعات المتساوية العدد المعادلة التالية :**

$$L.s.d = t_{(12,0.05)} \times \sqrt{\frac{2 \times Msr}{5}}$$

$$= 1.78 \times \sqrt{\frac{2 \times 2.26}{5}}$$

$$= 1.5$$

بعد ذلك نجد الفرق بين متوسطات الحسابية للمجموعات الثلاثة مع بعضها.

$$\bar{X}_1 - \bar{X}_2 = 0.67$$

$$\bar{X}_1 - \bar{X}_3 = 1.66$$

$$\bar{X}_2 - \bar{X}_3 = 2.33$$

القرار :

كلاً من المجموعة الأولى اصــغر من قيمة (L.s.d) وان المجموعتين الثانية والثالثة اكبر من قيمة (L.s.d) يدل على ان أنهما أفضل من المجموعة الأولى.

٤- اختبارات الفروض للبيانات المستقلة لعينين او اكثر كبيرة العدد

– اختبار مربع كاي (كا٢ – χ^2)

يعتبر اختبار (χ^2) واحداً من أكثر اختبارات الإحصاء أهمية ، لانه يستخدم للعديد من الأغراض، لهذا السبب سماه جيلفورد عام ١٩٥٦ إحصاء الغرض العام ، ويصنف هذا الاختبار في مجال الإحصــاء الاســتدلالي كاختبار بارومتري لأنه لا يشــمل على افتراضــات محددة فيما يتعلق باعتدالية توزيع البيانات او تجانس التباينات ويستخدم اختبار (χ^2) عندما تكون البيانات مأخوذة لعينات كبيرة مســـتقلة عندما يتم التعبير عن البيانات في شـــكل تكرارات (وحدات عد متكررة) او نسب او نسب مئوية ويطبق الاختبار فقط على البيانات المتقطعة.

ويستخدم اختبار (χ^2) بطرق متعددة لتحقيق بعض الأغراض الهامة هي :

اولاً : اختبار (χ^2) لعامل واحد

يستخدم (χ^2) لعامل واحد في حالات التوزيع متعدد الحدود في التجارب التي يتم فيها وصف مجموعة من المشاهدات عن طريق وضع كل مشاهدة في فئة واحدة فقط من بين فئات متعددة للتصنيف، حيث يستخدم التوزيع المتعدد الحدود لاختبار حسن المطابقة عن طريق التعرف على الوضع التقريبي لمربع كاي في هذا التوزيع.

ويمكن إيجاده وفق المعادلة التالية :

$$\chi^2 = \sum\left[\frac{(f_i - E_i)^2}{E_i}\right]$$

f_i : التكرارات E_i : القيمة المتوقعة

مثال : اشتركت مجموعة تتكون من ١٥٠ عداء في مسابقة للعدو ٢٠٠ م في مضمار لألعاب القوى من ست حارات ، وقد وزع المتسابقون على الحارات الست عشوائياً ، وأنهم قد التزموا بالأداء كل في حارته طوال مدة السباق ، وبعد الانتهاء من السباق حسبت التكرارات المشاهدة لعدد الفائزين بالمركز الأول من السباق لكل حارة من الحارات الست، كما حسبت التكرارات المتوقعة (النظرية) لكل حارة وفقاً لاحتمال النظري الذي يساوي ٠,١٦٧ ، فكانت ٠,١٦٧ × ١٥٠ =٢٥,٠٥ أي ٢٥ تقريباً لكل حارة من الحارات الست كالتالي:

الحارات (الخلايا)	التكرارات (المشاهدات)	التكرارات (المتوقعة)
١	٣١	٢٥
٢	٢٩	٢٥
٣	٢٢	٢٥
٤	٢٦	٢٥
٥	٢٠	٢٥
٦	٢٢	٢٥

فإذا كان احد الباحثين يرى ان الأداء في الحارة الداخلية (الحارة رقم (١)) تعطي ميزات خاصة للمتسابقين فيما تتمثل في أنها تمكنهم من مشاهدة منافسيهم أثناء بدء السباق مما يمنحه فرصة خاصة لتحقيق التفوق على منافسيه.

ولهذا الأسباب والتوقعات قام هذا الباحث بوضع الفرضي البحثي التالي :

في مسابقة العدو ٢٠٠م حيث يطلب من المتسابق الجري في حارته طوال مسافة السباق، فان الحارة الداخلية تمكن متسابقيها من الفوز نتيجة لما تقدمه لهم من مميزات خاصة.

وللتحقق من صحة هذا الفرض يتم حساب القيمة المشاهدة لمربع كاي– χ^2 ومقارنتها بقيمة χ^2 الجدولية وذلك وفق الخطوات التالية :

الخطوة : نقوم بحســــاب قيمة χ^2 في كل خلية من الخلايا (الحارات) بتطبيق معادلة χ^2 كالتالي :

الحارات (الخلايا)	التكرارات (المشاهدات) f_i	التكرارات (المتوقعة) E_i	$f_i - E$	$(f_i - E_i)^2$	$\left[\frac{(f_i - E_i)^2}{E_i}\right]$
١	31	25	6	36	1.44
٢	29	25	4	16	0.64
٣	22	25	−3	9	0.36
٤	26	25	1	1	0.04
٥	20	25	−5	25	1
٦	22	25	−3	9	0.36
المجموع	150	150	0		χ^2 =3.84

بما ان درجة الحرية تساوي عدد الخانات (الخلايا)-١ والتي تساوي ٦-١=٥ ان χ^2 الجدولية عند مستوى معنوية ٠,٠٥ تساوي 11.07 وهي اكبر من قيمة χ^2 المحسوبة =3.84 وهذا فأننا نرفض الفرضية البحثية (الفرضية البديلة) التي تقرر ان الحارة رقم ١ تمكن متسابقيها بالفوز بالسباق.

مثال : اراد إحدى الباحثين القيام باستطلاع رأي (60) طالباً في أحدى الجامعات حول أهمية مادة الإحصاء فأجاب (40)واحد منهم بنعم و(20) بلا ، المطلوب معرفة الفروق المعنوية للإجابات الطلبة حول أهمية مادة الإحصاء؟ علماً بان قيمة χ^2 = 3.84 عند مستوى معنوية 0.05 ودرجة حرية 1

الحارات (الخلايا)	التكرارات (المشاهدات) f_i	التكرارات (المتوقعة) E_i	$f_i - E$	$(f_i - E_i)^2$	$\left[\frac{(f_i - E_i)^2}{E_i}\right]$
١	40	30	10	100	3.33
٢	20	30	−10	100	3.33
المجموع	60	60	0		χ^2 =6.66

بما ان χ^2 الجدولية عند مستوى معنوية 0.05 تساوي 3.84 وهي اصغر من قيمة χ^2 المحسوبة =6.66 أي يوجد فروق معنوية في إجابة الطلبة نحو أهمية الإحصاء.

واجب : تم توجيه استبيان لـــ(140) طالب في أحدى الكليات عن ممارستهم للرياضية فأجاب (90) منهم بعدم الممارسة و (50) منهم يمارس الرياضة ، المطلوب هل هناك فروق معنوية بين الممارسين وغير الممارسين؟

ثانياً : اختبار (χ^2) لعاملين

في حالة وجود أكثر من متغيرين او أكثر ففي هذه الحالة نلجأ إلى أيجاد التكرار المتوقع بطريقة وأسلوب يختلف عما هو عليه في حالة متغير واحد ، ويمكن أيجاد التكرار المتوقع في حالة وجود متغيرين او أكثر وفق المعادلة التالية :

$$E = \frac{\sum_{i}^{n} f_i \times \sum_{j}^{n} f_j}{\sum_{i,j=1}^{n} f_{ij}}$$

$\sum_{i}^{n} f_i$: مجموع التكرارات الأفقية (الصفوف)

$\sum_{j}^{n} f_j$: مجموع التكرارات العمودية (الأعمدة)

$\sum_{i,j=1}^{n} f_{ij}$: المجموع الكلي للتكرارات

مثال : قام باحث باستطلاع رأي (100) طالب و(50) طالبة حول أهمية مادة الإحصاء فأجاب (30) بنعم و (70) بلا أما الطالبات فإجابة (10) منهن بنعم و(40) بلا.

المطلوب هل يوجد فروق معنوية في اجابة الطلبة وفقاً لمتغير الجنسين ؟ علما بان قيمة χ^2 تساوي 3.84

الجواب / الجنس	نعم	لا	المجموع
طالب	30a	70b	100
طالبة	10d	40c	50
المجموع	40	110	150

لايجاد القيمة المتوقعة لكل خانة نستخدم القانون التالي :

$$E = \frac{\sum_{i}^{n} f_i \times \sum_{j}^{n} f_j}{\sum_{i,j=1}^{n} f_{ij}}$$

$$E_A = \frac{\sum_i^n f_i \times \sum_j^n f_j}{\sum_{i,j=1}^n f_{ij}} = \frac{100\times 40}{150} = 26.66$$

$$E_b = \frac{\sum_i^n f_i \times \sum_j^n f_j}{\sum_{i,j=1}^n f_{ij}} = \frac{100\times 110}{150} = 73.33$$

$$E_C = \frac{\sum_i^n f_i \times \sum_j^n f_j}{\sum_{i,j=1}^n f_{ij}} = \frac{50\times 40}{150} = 13.33$$

$$E_d = \frac{\sum_i^n f_i \times \sum_j^n f_j}{\sum_{i,j=1}^n f_{ij}} = \frac{50\times 110}{150} = 36.66$$

نجد قيمة χ^2

الحارات (الخلايا)	التكرارات (المشاهدات) f_i	التكرارات (المتوقعة) $: E_i$	$f_i - E$	$(f_i - E_i)^2$	$\left[\frac{(f_i - E_i)^2}{E_i}\right]$
A	30	26.66	3.34	11.15	0.41
B	70	73.33	3.33	11.8	0.15
c	10	13.33	3.33	11.8	0.83
d	40	36.66	3.34	11.15	0.30
المجموع	200	200	0		χ^2 =1.69

درجة الحرية تساوي (عدد الصوف-١) + (عدد الاعمدة - ١) = ١

وبما ان χ^2 الجدولية عند مســتوى معنوية 0.05 تســاوي 3.84 واهي اكبر من قيمة χ^2 المحســوبة =1.69 لا يوجد فروق معنوية في إجابة الطلبة نحو أهمية الإحصـــاء نحو متغير الجنسين.

واجب : أراد إحدى الباحثين ان يقيم مدى استقلالية متغيري الجنس والاشتراك في الفرق الرياضية بالجامعة، فقام باستطلاع راي عينة عشوائية تتكون من 200 طالب من طلاب الجامعة ، وبعد تصنيفهم وفقاً لمتغيري الجنس وعضوية الفرق الرياضية بالجامعة حصل على البيانات الموضحة في الجدول التالي :

الجنس	مشترك	غير مشترك	المجموع
طالب	60	40	100
طالبة	40	60	100
المجموع	100	100	200

المطلوب : اختبار الفرضية الصفرية الذي يقرر ان الجنس والاشتراك في عضوية الفرق الرياضية بالجامعات متغيران مستقلان ، بمعنى ان الجنس عامل غير موثر (مستقل) بالنسبة للاشتراك في الفرق الرياضية بالجامعات، علماً بان قيمة χ^2 الجدولية تساوي 3.84 عند درجة حرية ١.

واجب : قام احد الباحثين بأجراء دراسة مسحية على عينة من طلبة احدى الكليات للتعرف على ارائهم في برنامج النشاط الرياضي الداخلي الذي تنظمه الكلية وكان الباحث يرغب في تحديد ما اذا كانت العينة التي يستخدمها تمثل المجتمع الطلابي للكلية ام لا؟ وكان حجم العينة التي يستخدمها الباحث يضم (400) طالب موزعة على السنوات الدراسية الاربع كالتالي:

السنة الدراسية	التكرارات (المشاهدات) f_i	النسبة المئوية	التكرارات (المتوقعة) E_i :
الاولى	50	30%	120
الثانية	100	20%	80
الثالثة	150	25%	100
الرابعة	100	25%	100
المجموع	400	100%	400

المطلوب معرفة اذا كانت العينة تمثل مجتمع كلية التربية الرياضية ام لا

الحل :

السنة الدراسية	التكرارات (المشاهدات) f_i	النسبة المئوية	التكرارات (المتوقعة) $: E_i$	$f_i - E$	$(f_i - E_i)^2$	$\left[\frac{(f_i - E_i)^2}{E_i}\right]$
الاولى	50	30%	120	-70	4900	40.8
الثانية	100	20%	80	20	400	5.0
الثالثة	150	25%	100	50	2500	25.0
الرابعة	100	25%	100	0	0	0
	400	100%	400			χ^2 =70.8

بما ان درجة الحرية تساوي عدد الخانات (الخلايا)-١ والتي تساوي ٦-١=٥ ان χ^2 الجدولية عند مستوى معنوية ٠,٠٥ تساوي 11.34 واهي اصغر من قيمة χ^2 المحسوبة =70.8 وهذا المجتمع اخذ بطريقة التحيز وليس موزع توزيع طبيعي.