وزارة التعليم العالي والبحث العلمي

جامعة تكريت

كلية علوم الأغذية الشرقاط

قسم تكنلوجيا الألبان

المرحلة الأولى

محاضرات في الإحصاء الحياتي
لقسم تكنلوجيا الألبان - المرحلة الأولى

للمدرس المساعد: خيرالله فرج سبهان الجبوري

الفصل الأول

- تعريف الإحصاء الحياتي
- الإحصاء الوصفي.
- الإحصاء الاستدلالي.
- العلاقة بين القياس والإحصاء الحياتي.
- المجتمع .
- العينة.

نبذة عن علم الإحصاء

يعد الإحصاء من الأمور المعروفة لدى المجتمعات القديمة، فقد استخدمت كلمة إحصاء ((Statistic لأول مرة في مجال الشؤون المتعلقة بإعمال الدول والحكومات، وخاصة تلك المتعلقة بقضايا التنظيم وإدارة شئون البلاد، وجمع الضرائب وما شابه. وكما استخدم الإحصاء في الأغراض العسكرية والحربية واستعراض القوة، وذلك من خلال حصر عدد الجنود والأسلحة. ثم توسعت عمليات التعداد والحصر لتشمل بيانات عن المواليد والوفيات والإنتاج والاستهلاك . وبذلك نشأت الحاجة، ثم توسعت عمليات التعداد والحصر على إحصاء عدد أفرادها من أجل معرفة قوتها وكثرتها، لتشمل بيانات عن المواليد والوفيات والإنتاج والاستهلاك. وبذلك نشأت الحاجة إلى تنظيم وتلخيص هذه البيانات ووضعها في صورة او رسم بياني حتى يسهل الرجوع إليها والاستفادة منها بأسرع وقت ممكن، وقد أطلق على هذه الطرق علم الدولة أو علم الملوك ثم علم الإحصاء.

وقد استخدم العرب الإحصاء في العصور الإسلامية المبكرة. وقد وردت كلمة الإحصاء ومشتقاتها في القرآن الكريم إحدى عشرة مرة، منها قوله تعالى **(وَكُلَّ شَيْءٍ أَحصَيْناهُ كِتابا** ,النبأ: ٢٩)

وقد نشأ علم الإحصاء في أطار النظام السياسي للدولة على يد البارون فون بيفيلد Von Biefeld, ١٧٧٠. وفي أواخر القرن التاسع عشر طورت أساليب وأفكار إحصائية، حيث ترجع النشأة الرياضية الصحيحة لهذا العلم إلى أبحاث الرياضي الفرنسي لابلاس (Laplace) والرياضي الألماني جاوس Gaus والعالم الأنكليزي فرانسيس غالتون Galton والرياضي الأنكليزي كارل بيرسون (Karel Pearson).

هذا كل ما كان يعرف عن علم الإحصاء في ذلك الوقت، حيث كان التحليل الإحصائي يستخدم للوصول إلى نتائج يستفاد بها في اتخاذ القرارات من الأشياء التي لم تستخدم بعد رغم أنه قد ظهرت الحاجة الماسة لاستخدامها ولا سيما بعد تطور علم الاحتمالات في القرنين السابع عشر والثامن عشر الميلاديين. هذا وقد ظل الاعتقاد في ذلك الوقت بأن علم الإحصاء هو العلم الذي يختص بالطرق العلمية لجمع وتنظيم وعرض البيانات، إما في صورة بيانية أو صورة جدوليه، حتى أن بعض الأشخاص قليلي الاطلاع ومحدودي التعليم في الوقت الحاضر يعتقدون أن الإحصاء ما هي إلا هذه الطرق. إلا أنه بعد التطور أصبحت الحاجة ملحة إلى تحليل البيانات التي جمعت كالتنبؤ بعدد السكان بعد فترة زمنية بناء على التعدادات الموجودة أو التنبؤ بالإنتاج والاستهلاك، أو طرق أخذ العينات وتصميم التجارب. وقد ساعد على ذلك تطور علم الاحتمالات الذي كان له دور كبير في تحليل البيانات واتخاذ القرارات المناسبة بناء على هذا التحليل.

مفهوم الإحصاء :

الإحصاء كمصطلح أشتق من عدة كلمات قديمة، فهو مأخوذ من الكلمة اللاتينية (Status)، وعن الألمانية (Stutstistic)، والايطالية (Stator)، وكلها تعني سياسة شؤون الدولة. وورد مصطلح الإحصاء في قاموس ويبستر (Webestars, 1971) كحقائق مصنفة تمثل معلومات عن الفرد في الدولة بحيث يمكن وصفها بإعداد أو أي صيغة كمية، سواء كانت تبويبا أو تطبيق.

وعلم الإحصاء (Statistics) هو أحد فروع الرياضيات الهامة ذات التطبيقات الواسعة. وهو العلم الذي يهتم بجمع المعلومات والبيانات وتنظيمها وتحليلها وتلخيصها وعرضها لتحديد اتجاهات تطور الظواهر الطبيعية والإنسانية، وأنه يشمل على مجموعة من الطرق الإحصائية Statistical Methods التي تستخدم في اتجاهات مختلفة. ويعرف علم الإحصاء بأنه ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع وتنظيم وتلخيص وعرض وتحليل البيانات، وذلك للوصول إلى نتائج مقبولة وقرارات سليمة على ضوء هذا التحليل.

اذاً, يهتم علم الإحصاء بجمع وتلخيص وتمثيل وإيجاد استنتاجات من مجموعة البيانات المتوفرة، محاولا التغلب على مشاكل مثل عدم تجانس البيانات وتباعدها .كل هذا يجعله ذا أهمية تطبيقية واسعة في شتى مجالات العلوم من الفيزياء إلى العلوم الاجتماعية وحتى الإنسانية، كما يلعب دورا في السياسة والأعمال.

وكلمة الإحصاء بالنسبة للباحثين في العلوم التربوية والنفسية تعني الطرق والإجراءات التي يستخدمها الباحث في محاولته لفهم بيانات عن ظاهرة ما .وبالرغم من ان الإحصاء يعد من العوامل المهمة التي ساهمت في تطوير البحوث التربوية والنفسية، ولكن أصبح تعلمه ضرورة ليس في مجال البحث العلمي فحسب، وإنما لكل العاملين في مجال التربية وعلم النفس .كالمعلم والإداري والمشرف والمخطط التربوي والباحث الاجتماعي والمرشد التربوي وغيرهم .فكل هؤلاء بحاجة الى المعرفة الأساسية في الإحصاء لأسباب شتى، إذ ان اغلب المقالات والكتب والبحوث والدراسات العلمية ذات العلاقة المباشرة بأعمال التربويين والنفسانيين التي تنشر يوميا لا تخلو من البيانات

تعريف الاحصاء الحياتي : هو رياضيات لدراسة الظواهر البايلوجية من خلال جمع البيانات وتلخيصها وتحليلها وتعميم نتائجها- مثل الدراسة الاحصائية لفصائل الحيوان والنبات ومعرفة خواص كل جنس بينهما

فروع الإحصاء الحياتي

ينقسم علم الإحصاء إلى فرعين رئيسين:

أولا: الإحصاء الرياضي:Mathematical Statistics ويتناول اكتشاف أو استنتاج القوانين والنظريات الإحصائية، وفقا لأسس رياضية.

ثانياً: الإحصاء التطبيقي Applied Statistics: يستخدم الإحصاء التطبيقي النظريات والقوانين

الإحصائية في عمليات التنظيم والتحليل والمقارنة والاستنتاج في البحوث العلمية، التي تجرى في شتى الميادين. ويمكن تقسيم الإحصاء التطبيقي من الناحية الوظيفية الى فرعين أساسيين هما:

١- **الإحصاء الوصفي Descriptive Statistics:** وهو الإحصاء الذي يتناول تنظيم وعرض ووصف البيانات والمعلومات سواء كانت هذه المعلومات كمية، كالوزن والعمر والطول. أو كانت نوعية، كالجنس والحالة الاجتماعية وأنماط الشخصية. فهو بذلك يعتبر أداة يستعين بها الباحث لتلخيص أو تركيز مجموعة من البيانات، وينظمها بشكل يسهل فهمه واستيعابه. ويستخدم هذا الفرع من الإحصاء مختلف الوسائل الوصفية: البيانات المفردة والمبوبة كالجداول والتوزيعات التكرارية، والرسوم البيانية، ومقاييس النزعة المركزية، ومقاييس التشتت، ومقاييس العلاقة، ومختلف القياسات الأخرى. حيث تتضمن أساليب الإحصاء الوصفي وصف وتوزيع مجموعة كبيرة من القيم الخاصة بأحد المتغيرات بحسب أحد مقاييس النزعة المركزية كالمتوسط الحسابي او الوسيط او المنوال او حساب أحد مقاييس التشتت كالانحراف المعياري او نصف المدى الربيعي ولا تقتصر تلك الأساليب على دراسة متغير واحد بل تتعداها الى وصف العلاقة بين متغيرين او أكثر وصفا علميا حقيقيا بحساب أحد معاملات الارتباط. ومن ذلك يتسنى للباحث ان يتنبأ بقيم أحد المتغيرين بفضل ما يعرفه عن المتغيرات الأخرى التي ترتبط بذلك المتغير باستعمال مفهوم الانحدار.

٢- **الإحصاء الاستدلالي Inferential Statistics:** يقصد بالإحصاء الاستدلالي تلك العملية التي تؤدي الى استخلاص النتائج العامة من النتائج الجزئية على وفق قوانين إحصائية معينة. إذا، يركز هذا النوع من الإحصاء على الوصول إلى استنتاجات حول خصائص المجتمع من خلال استخدام المعلومات المتوفرة عن العينة المحسوبة من هذا المجتمع، أي انه يهدف الى التعميم من العينة إلى المجتمع. وبهذا فان هذا الإحصاء يركز على اختبار الفرضيات المتعلقة بالفروض بين المتوسطات أو النسب المؤية المتعلقة بعينة واحدة أو عينتين أو أكثر أو الفروق بين معاملات الارتباط.

ونظرا لعدم إمكانية دراسة جميع أفراد المجتمع موضوع البحث لما يتطلبه ذلك من قدر كبير من الوقت والجهد والمال فقد أصبح الاعتماد على أسلوب العينات ضروريا خاصة علم الإحصاء قد بلغ من التقدم مرحلة يستطيع مهما يستنتج ما يريد استنتاجه عن المجتمع الذي اختيرت منه هذه العينة بدرجة عالية من الثقة وفي حدود محتملة من الخطأ.

حيث تتضمن أساليب الإحصاء الاستدلالي تقييم مؤشرات المجتمع من القيم المناظرة لها والمحسوبة من بيانات العينة التي يطلق على كل منها إحصاء. قد يكون التقدير في صورة قيمة واحدة او فترة من القيم تسمى بفترة الثقة.

أهمية علم الإحصاء

يعد علم الإحصاء أحد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع البيانات والمعلومات اللازمة في البحث العلمي وتحليل هذه المعلومات لغرض الوصول إلى النتائج التي يهدف لها البحث كما وان للإحصاء دورا بارزا في وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج لكافة القطاعات حيث يمكن تطبيق علم الإحصاء في مجالات العلوم الإنسانية أو العلوم الصرفة (مثل الزراعة، الصناعة، الطب، الهندسة، الإدارة والاقتصاد، وغيرها).

ولقد كان للإحصاء دور بارز في تقدم علم النفس وأساليب القياس المستعملة فيه وبخاصة تلك المتعلقة بعمليات القياس العقلي والخصائص النفسية والتحليل العاملي كأحد الفروع التطبيقية لعلم الإحصاء ظهر في مجال القياس العقلي واستعد أسسه من الحقائق الرياضية المتمثلة في جبر المصفوفات وهندسة والتحليل المتجهات والتفاضل كما أدى استعمال الإحصاء في علم النفس الى تطور البحوث النفسية التي تهدئ الى دراسة البيئة والوراثة والتفاعل بينهما والأثر الناتج عن هذا التفاعل في سمات الشخصية والسلوك البشري.

ويعد علم الإحصاء بحد ذاته وسيلة وليس غاية مما يعني استخدامه أينما وجد في البحث العلمي، وتنطوي أهمية علم الإحصاء في الحياة العملية على ما يأتي:

1. يعد علم الإحصاء أحد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في عملية:
   - جمع وتلخيص البيانات والحقائق أو المشاهدات حول الظواهر.
   - عرض هذه البيانات بصورة واضحة ومحددة.
   - تحليل البيانات. ووضع أسس مقارنة المتغيرات الخاصة بالظاهرة قيد الدراسة.
   - اختبار الفروض الإحصائية وتفسير النتائج.
2. للإحصاء دور بارز في التنبؤ المستقبل ووضع الخطط المستقبلية عن طريق التنبؤ بالظاهرة من خلال النتائج.
3. وضع الخطط واتخاذ القرارات المطلوبة بصددها وذلك لما يتوفر من بيانات ذات علاقة.

وأما بالنسبة للمعلم فان علم الإحصاء يساعده في:

- عرض نتائج الطلاب بشكل بيانات مبوبة.
- التعرف على نسبة النجاح والرسوب.
- ترتيب الطلاب حسب درجاتهم.
- إجراء البحوث التربوية والنفسية، وفهم وتفسير نتائجها.

العلاقة بين القياس Ăلÿلاحصاء

مر علينا ان علم الإحصاء يتناول طرق جمع البيانات الخاصة بمختلف الظواهر وعرضها وتحليلها للوصول إلى نتائج تساعد في اتخاذ القرارات المناسبة، وذلك من خلال فرعين: الإحصاء الرياضي والإحصاء التطبيقي، والذي بدوره يقسم الى الإحصاء الوصفي والإحصاء الاستدلال.

أما القياس فهو نظام تصنيفي يتم بواسطته التعبير عن صفات الأشياء والحوادث بأعداد وأرقام او رموز (الوصف الكمي)، ويكون هذا التكميم حسب قواعد محددة ودقيقة. فلأية ظاهرة خصائص وسمات كثيرة، تميزها عن غيرها من الظواهر، فعندما يحاول الباحث القيام بدراسة وملاحظة هذه الخصائص والسمات المميزة للظاهرة، فأنه يحتاج إلى استخدام نظام معين لتصنيف وتسجيل هذه الخصائص والسمات طبقاً لقواعد مصاغة صياغة محددة وواضحة. إذا إن عملية إعطاء أرقام خاصة للأشياء المختلفة، هي ما يسمى عادةً بالقياس، فالقياس هو نظام تصنيفي تعطى فيه للأشياء أرقام خاصة بها، وبذلك فان القياس يشمل جمع البيانات والملاحظات الكمية (الرقمية) عن الصفة او السمة المراد قياسها. لكي يسهل تسجيلها وتلخيصها ومعالجتها وتحليلها باستعمال الأساليب الإحصائية.

وعملية القياس تتم من خلال نوعين (الطبيعي الفيزيائي المباشر)، و(التربوي والنفسي غير المباشر). فالقياس الفيزيائي يشمل تحديد سمات الطول، والعرض، والارتفاع، والوزن والحجم، والحرارة، والصوت، والضوء، وكلها ظواهر فيزيائية تحتاج الى تحديد بمقاييس فيزيائية مختلفة، وهذا يتسم بالدقة. أما القياس التربوي فيشمل قياس سمات التحصيل الصفي لدى الطلبة ومدى تفاوت درجاتهم وأضيف الى هذا النمط من القياس، الناحية النفسية كقياس مستوى القلق والذكاء، ومفهوم الذات.

المجتمع الإحصائي والعينة الإحصائية

المجتمع الإحصائي هو عبارة عن أي تجمع معرف من الأشياء أو الأشخاص أو الحوادث. ويعرف المجتمع على أنه مجموعة من الأفراد محل الدراسة والتي لها خصائص مشتركة. وهو المجموعة الشاملة التي يجرى اختيار العينات منها. اذاً، المجتمع الإحصائي يشمل جميع المفردات موضع الدراسة والتي نرغب في معرفة حقائق عنها سواء كانت على شكل إنسان أو حيوان أو جماد أو درجات امتحان أو منازل أو مزارع أو سفن ... الخ.

ويقسم المجتمع الإحصائي إلى قسمين:

1. **مجتمع محدود (Finite): والذي** يكون فيه عدد محدود من المفردات او الأفراد مثل تلاميذ الصف الأول الابتدائي في منطقة معينة، أو عدد تلاميذ الصف الثاني الابتدائي ضعيفي القراءة في اللغة العربية. أو قد يكون إنتاج شركة من دواء معين.,...إلخ.
2. **مجتمع غير محدود** (Infinite): هو المجتمع الذي يكون فيه عدد الأفراد غير منته مثل عدد النجوم في السماء، عدد حبات القمح المحصود في مزرعة معينة...إلخ.

أساليب جمع البيانات:

أولاً: أسلوب التسجيل الشامل Census: ويقصد به جمع بيانات عن جميع المفردات التي تؤلف المجتمع الإحصائي هنا يجب أن يكون المجتمع محدد.

ثانياً: أسلوب العينات Samples: ويقصد به جمع البيانات والمعلومات عن جزء من المفردات التي تؤلف المجتمع الإحصائي. وهو أسلوب مفيد في المجتمعات الغير محدودة كما وتحتاج إلى وقت وجهد وموارد مادية وبشرية اقل مما يحتاجه أسلوب التسجيل الشامل.

وفى بعض الحالات لا نتمكن من حصر كل المفردات للمجتمع مثل مجتمعات الأسماك أو النباتات، أو أن تؤدي عملية الحصول على البيانات لمفردات المجتمع إلى إتلافها أو هلاكها، مثال ذلك فحص دم المريض كله يؤدي إلى وفاة الشخص، وبالتالي يتم جمع البيانات عن جزء فقط من مفردات المجتمع يسمى العينة (Sample).

العينة الإحصائية Statistical sample.

العينة الإحصائية هي جزء من مفردات المجتمع يتم اختيارها بحيث تكون ممثلة للمجتمع ككل. وأسلوب أخذ العينات شائع الاستعمال عند إجراء الدراسات والبحوث الإحصائية لأن تكاليفه أقل، وبواسطته يمكن الحصول على نتائج سريعة، مقارنة بأسلوب الحصر الشامل الذي يتم فيه جمع البيانات من كل مفردات المجتمع. وتمثل العينة على سبيل المثال جزء من سكان مدينة معينة أو جزء من منازل هذه المدينة أو جزء من درجات التلاميذ لأحد المقررات الدراسية وهكذا. وتكون العينة في الغالب محدودة في عدد أفرادها، ويتم التحكم في طريقة اختيار أفراد العينة وفي عددهم من خلال بعض المبادئ الإحصائية ويوجد علم خاص بطرق أخذ العينات يسمى المعاينة الإحصائية (Statistical Sampling).

ومن العينة الإحصائية يتم الوصول إلى نتائج يمكن تعميمها على المجتمع الإحصائي محل الدراسة ككل. وفى حالة احتمال عدم تمثيل العينة تمثيلا حقيقيا، فإن الاستدلال الإحصائي يمكن الباحث من قياس الخطأ الناتج عن ذلك. وهناك طرق كثيرة لتحديد كيفية أخذ العينة الممثلة للمجتمع.

أنواع العينات.

تقسم العينات إلى قسمين: الاحتمالية او العشوائية، ألا احتمالية او العمدية:

القسم الأول: المعاينة الاحتمالية: Probability Sample

يعتمد هذا الأسلوب على نظرية الاحتمالات في اختيار العينة بطريقة عشوائية أي غير متحيزة حيث يكون لإفراد المجتمع فرص متساوية لاختيارهم ضمن العينة، ولا يكون للباحث دخل أو رأي في اختيار هذه المفردة أو تلك وإنما على أساس عشوائي. وهنا يجب تحديد مجتمع الدراسة بكل دقة ومن ثم الانطلاق لانتقاء العينة. وفيما يلي شرح موجز لأنواع العينات العشوائية:

١- **العينة العشوائية البسيطة Simple Random Sample:** وهي أسهل أساليب المعاينة العشوائية وابسطها، يتم اللجوء إليه فقط في حالة توفر شرطين أساسيين وهما: أن تكون كافة أفراد مجتمع البحث معروفين، وحينما يكون هنالك تجانس بين هؤلاء الأفراد. وهي العينة التي يتم تكوين عناصرها على أساس تساوي الفرص أمام جميع المجتمع الإحصائي للظهور في العينة أو ان أية مجموعة جزئية من المجتمع الإحصائي وبحجم معين لها الفرصة نفسها لتختار كعينة من ذلك المجتمع ويتم ذلك عمليا بطريقتين: طريقة القرعة، أو طريقة جدول الأرقام العشوائية

٢- **العينة العشوائية المنتظمة Systematic sample:** شرط سحبها هو أن يكون المجتمع الإحصائي مرتبا ترتيبا تصاعديا أو تنازليا أو وفق أي ترتيب أخر، حيث نقسم المجتمع المرتب إلى (L) طبقة. ثم يتم سحب عينة عشوائية بسيطة من الطبقة الأولى فقط ثم نضيف (K) لنسحب المفردة الثانية وهكذا بحيث نضمن سحب مفردة من كل طبقة. وهكذا، تعتمد هذه العينة على اختيار الفرد بطريقة عشوائية ثم يختار عدد أفراد العينة المطلوبة فيتم اختيار الأفراد ذوي التسلسل المنتظم، ويفضل استخدام هذا النوع من العينات في حالة تجانس المجتمع. أي عندما تكون مفردات المجتمع منتظمة أو دورية تتسم بعدم التغير أو بالتباين الشديد في مفرداتها وتمتاز هذه الطريقة بانتظام وثبات التباعد بين وحدات العينة. وفي الغالب يتم إختيار العينة المنتظمة عن طريق حصر كافة مفردات مجتمع الدراسة الأصلي ومن ثم إعطاء كل فرد رقم متسلسل ثم قسمة عدد مفردات مجتمع البحث على حجم العينة المطلوبة مما ينتج الرقم الذي سيفصل ما بين كل مفردة يتم اختيارها في عينة الدراسة والمفردة التي تليها. وفي الغالب إختيار المفردة الأولى يكون عشوائي.

٣- **العينة العشوائية الطبقية Stratified Random Sample:** شرط سحبها أن يكون المجتمع غير متجانس. فإذا وجد الباحث أن المجتمع الإحصائي غير متجانس يتم تقسيمه لعدة أقسام حسب متغيرات الدراسة كالجنس أو المستوى الاقتصادي أو الاجتماعي، أو التخصص في الدراسة (علمي، أدبي، مهني)، بحيث نستطيع تحويله إلى مجتمع متجانس، في هذه الحالة فانه من الممكن ان يلجا الباحث إلى اعتبار كل طبقة وحدة واحدة. ومن ثم يتم سحب عينة عشوائية بسيطة من كل طبقة بحيث يتناسب حجمها مع حجم الطبقة في المجتمع، بحيث تكون نسبة الأفراد من طبقة معينة في هذه العينة فيما بين الطبقات وخاصة إذا كان المتغير الذي قسم المجتمع إلى طبقات في ضوئه ذا علاقة بالمتغير التابع.

٤- **العينة العشوائية ذات المراحل المتعددة Multi-stage random sample:** شرط سحبها هو أن يكون المجتمع الإحصائي كبير جدا وموزع على بقعة جغرافية واسعة كأن يكون المجتمع على مستوى البلد أو المحافظة أو المقاطعة أو إقليم هنا نبدأ بتقسيم المجتمع إلى وحدات أولية ثم نأخذ عينة عشوائية من هذه الوحدات كمرحلة أولى. ثم نقسم كل وحدة أولية مختارة إلى وحدات أصغر (الوحدات الثانوية) ويتم اختيار عينة من الوحدات الثانوية لكل وحدة أولية كمرحلة ثانية. ونستمر بالتقسيم والاختيار على هذا المنوال إلى إن نصل إلى عدد المفردات التي تؤلف العينة العشوائية

القسم الثاني: المعاينة اللا احتمالية Probability Sample - Non: هنا يتدخل الباحث شخصيا في اختيار مفردات العينة وبصورة متحيزة وليس على أساس عشوائي لاعتقاده بأن هذه المفردات هي خير ما تمثل المجتمع والتي ستغني الدراسة بآرائها وأطروحاتها، إذ يختار عينة حسب مواصفات ومعايير يضعها او يختار كيفما اتفق ممن يجدهم في مكان. أو من ألمعارف أو من المتطوعين، أو من إفراد الذين عاصروا ظاهرة أو قضية معينة. وبصورة عامة فان العينات غير الاحتمالية قد تكون مفيدة من ناحية تقليل الجهد والتكاليف إلا انه يصعب تعميم نتائجها على جميع أفراد المجتمع الإحصائي. واهم أنواع هذه العينات هي:

١- **العينة العرضية Accidental Sample:** وفي هذا النوع يتم اختيار العينة عن طريق الصدفة، كأن يوزع الباحث استبانة على طلبة الجامعة من يشاهدون إحدى المسرحيات أو ممن هم جالسون في مكان ما ليأخذ رأيهم في قضية عامة في الجامعة. وهذا النوع من العينات يصعب تعميم نتائجها على المجتمع، إضافة إلى أنها قد تكون مضللة.

٢- **العينة ألعمدية Purposive sampling:** وهي العينة المختارة بشكل متعمد فنعتقد مسبقا ان مفردات هذه العينة هي خير من يمثل مجتمع الدراسة. فمثلا عند دراسة السبل الكفيلة للارتقاء برياضة كرة القدم فمن الأفضل اختيار عينة من المتخصصين برياضة كرة القدم وبشكل عمدي كون ان هذه العينة هي ذات خبرة بشؤون هذه الرياضة.

يعتمد الباحث هذه العينة للتخلص من المتغيرات الدخيلة مثل اختيار مجموعة من المعلمين لأخذ رأيهم في العملية التربوية. وتستخدم في دراسات استطلاع الرأي العام، مثل ما تقوم به وسائل الإعلام من استطلاع آراء أولياء الأمور والخبراء حول النظم المقترحة لتطوير الدراسة بالثانوية العامة مثلاً.

٣- **عينة الحصص Quota sampling**: بموجب هذا النوع من المعاينة يتم تقسيم المجتمع الإحصائي الى عدة طبقات بالاستناد الى معايير تتعلق بطبيعة الدراسة، ثم يتم اختيار عينة عمدية من كل طبقة (بشكل عشوائي) يتناسب حجمها وحجم الطبقة في المجتمع، ومجموع أحجام هذه العينات ألعمدية يمثل حجم عينة الحصص.

الرموز الإحصائية:

الاسم	الرمز	الاسم	الرمز
الفئات	C	معامل الاختلاف	C.V
التكرار	F	قيمة مطلقة	| |
التكرار النسبي	P	معامل الارتباط	r
المدى	R	معامل بيرسون للارتباط الخطي	r_p
تكرار الفئات	K	معامل سبيرمان للارتباط الرتب	r_s
طول الفئة	L	الانحدار	$\tilde{y}$
الحد الحقيقي للفئة	A	درجة خام لمتغير عشوائي	X
التكرار المتجمع الصاعد	f↓	درجة معيارية	Z
الوسط الحسابي (مجتمع)	μ	درجة خام لمتغير عشوائي ثاني	y
الوسط الحسابي (عينة)	$\ddot{X}$	مجموع	$\sum$
الوسيط	Me	الخطاء المعياري	S.E
عدد المفردات	N او n	معامل الالتواء	S.K
المنوال	M0	معامل التفرطح	K
المتوسط المرجح	$\ddot{X}_w$	درجة الثقة	1- α
التباين (مجتمع)	σ^2	مستوى المعنوية	α
التباين (عينة)	S^2	فرض العدم	H_0
الانحراف المعياري (مجتمع)	σ	فرض البديل	H_1
الانحراف المعياري (عينة)	S		

الفصل الثانـي

- طرائق عرض البيانات.
- الأعمدة البيانية.
- المدرج التكراري.
- المضلع التكراري.
- الخط المنحنـي التكراري.
- الدائرة البيانية.

طرق عرض البيانات:

البيانات الإحصائية عبارة عن معلومات صحيحة ودقيقة تجمع من مصادر محددة، وبطريقة سليمة. وهذه البيانات إما تكون: كمية (رقمية)، أو كيفية (وصفية). والخطوة التالية بعد جمع البيانات في مجال الإحصاء الوصفي، هو تبويب البيانات وعرضها بصورة يمكن الاستفادة منها في وصف طبيعة الظاهرة محل الدراسة، من حيث تمركز البيانات، ودرجة تجانسها. وان عرض البيانات في العديد من المجالات من أهم الأمور، وتخضع البيانات على اختلافها في العادة إلى العديد من العمليات التي تتم عليها والتي تعمل على معالجتها وتحويلها إلى قراءات يمكن توظيفها في المجال الذي يهتم بها، فهناك طرق تشجع المتلقي على التفاعل بشكل أكبر من الطرق الأخرى؛ بسبب سهولة عرض البيانات، وسهولة استيعابها، وهناك نوعين للبيانات هما:

أولاً: البيانات النوعية: وهي البيانات التي لا يمكن التعبير عن مفرداتها بأرقام عددية مثل الصفات، كنوع الجنس (ذكور، إناث)، وفرع الدراسة (علمي، أدبي)، والحالة الاجتماعية (متزوج، أعزب، مطلق، أرمل). وكذلك التقدير في الامتحان (ضعيف، مقبول، جيد، جيد جداً، ممتاز)، ولون الشعر، والجنسية، وهكذا بقية المتغيرات النوعية.

ثانياً: البيانات الكمية (العددية): أما البيانات الكمية يمكن التعبير عن مفرداتها بقيم عددية مثل درجة الطالب في الامتحان، أو الدخل، او الطول، او الوزن، او العمر، وهكذا بقية المتغيرات الكمية. والبيانات الكمية على نوعين:

أ- المستمرة او المتصلة: وهي البيانات التي تصف كمياً مجموعة الخصائص والمتغيرات التي يمكن لها ان تأخذ أي قيمة محتملة ضمن مدى او مجال معين. ويمكن استخدام (أجزاء الدرجة) كالكسور العشرية في الإشارة إلى مقدار الصفة كالطول، والوزن، والعمر، ودرجة الحرارة ويمكن أن تأخذ أي قيمة أي لا تحوي قفزات فالترمومتر يرتفع أو ينخفض ماراً بكل القيم.

ب- المتقطعة او المنفصلة: وهذه البيانات تأخذ قيماً محدودة ضمن المدى، ولا يمكن التعبير عنها بالكسور، ومنها ما يسمى بالمتغيرات الوثابة، وهي المتغيرات المرتبطة بالأعداد (عدد الطلاب، عدد الأقلام، عدد الشعب، عدد المدارس).

ويتم عرض البيانات الإحصائية بعدة طرق: (التوزيعات التكرارية، والأعمدة البيانية، والمدرج التكراري، والمضلع التكراري، والمنحنى التكراري، القطاعات الدائرية).

١- **التوزيعات التكرارية:** وسيلة لتلخيص وتنظيم البيانات في عدد محدود من الفئات في صورة جداول أو رسوم بيانية، تسهل طريقة فهمها بمجرد النظر إليها سواء كانت بيانات وصفية او كمية، من خلال ترتيبها تنازليا او تصاعديا وحساب مرات تكرارها. وتشمل التوزيعات التكرارية:

التوزيع التكراري البسيط: يستخدم عندما يكون حجم البيانات صغير، على سبيل المثال: إذا أردنا تكوين جدول تكراري بسيط بدرجات (25) تلميذ في اختبار تحصيلي ما وكانت الدرجات هي (8, 7, 8, 9, 7, 6, 5, 5, 8, 10, 6, 6, 8, 9, 7, 10, 7, 9, 9, 10, 7, 7, 8, 6). علينا إتباع الخطوات الآتية: ١. نرتب الدرجات (المشاهدات) ترتيبا تصاعديا او تنازليا. ٢. ثم نبحث عن أصغر الدرجات وهي (5)، ونبحث عن أكبر الدرجات وهي (10). ٤. نصمم جدول وندرج الدرجات في عمودين او صفين الأول يمثل الدرجات والثاني يمثل علامات تكرار كل درجة، والثالث يمثل تكرار كل درجة. كما في الجدول الأتي.

الدرجة (x)	علامات التكرار	التكرارات (f)
5	//	2
6	////	4
7	/ ///	6
8	///	5
9	////	4
10	///	3
المجموع		24

التوزيعات التكرارية للبيانات المبوبة (فئات الدرجات): عندما يزداد الفرق بين أكبر درجة وأصغر درجة، فإننا نستغرق وقت وجهد لإعداد جدول لتوزيع الدرجات وتسجيلها في صورة واضحة، ولهذا تجمع الدرجات في فئات ويكون علينا حساب مرات تكرار درجات كل فئة، وكل ذلك يتطلب معرفة المدى الكلي للدرجات، وتقسيم هذا المدى الى عدد من الفئات متساوية الطول وذلك بإتباع الآتي:

١. نحدد عدد الدرجات (n) وهم عدد التلاميذ، ثم نحدد أكبر الدرجات وأصغرها.

٢. نحسب المدى الكلي من المعادلة (المدى الكلى = أكبر درجة – أصغر درجة + 1).

٣. نحدد عدد الفئات المطلوب في ضوء طول الفئة من العلاقة: (عدد الفئات= المدى الكلي على مدى الفئة).

٥. نحدد بداية الفئة الأولى بأصغر درجة ويضاف إليها مدى الفئة لنحصل على نهاية الفئة الأولى. وتبدأ الفئة الثانية حيث انتهت الفئة الأولى ثم يضاف إليه مدى الفئة لنحصل على نهاية الفئة الثانية، وهكذا حتى نحصل على آخر الفئات.

٦. يحسب مرات تكرار كل درجة داخل كل فئة ويوضع أمامها.

مثال: إذا كانت الآتية تمثل درجات (30) تلميذ في اختبار تحصيلي:

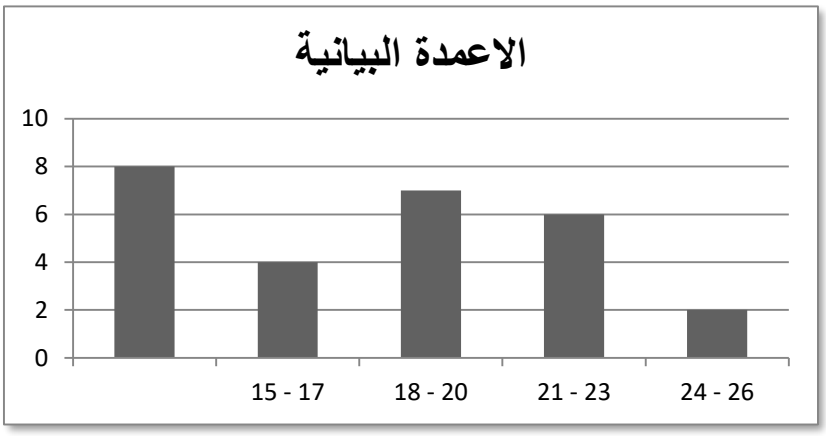
(٢١ ١٧ ١٣ **١٢** ٢٧ ١٣ ١٨ ٢٢ ١٨ ٢٧ ٢٢ ٢٠ ١٤ ١٦ ١٨ ٢٠ ٢١
٢٧ ١٣ ٢٢ ٢٣ ٢٠ ٢١ ١٢ ١٨ ١٧ ١٦ ١٤ ٢٥ ٢٦)

ولعرض هذه الدرجات في جدول تكراري نتبع الخطوات السابقة، لتصميم الجدول التكراري الآتي.

الفئات	علامات التكرار	التكرارات (f)
14 -12	/// ///	8
17 -15	////	4
20 -18	// ///	7
23 -21	/ ///	6
26 -24	//	2
29 -27	///	3
المجموع		30

٢- الأعمدة البيانية. وتسمى بالأعمدة الرأسية، وهي عبارة عن أعمدة او مستطيلات ذات قواعد صغيرة متساوية غير متلاصقة، على المحور السيني (x)، وترتفع بأطوال متباينة تبعا لتباين مستويات المتغير التي يمثلها على المحور الصادي (y). وتتميز هذه الأعمدة بكونها أشكالا بسيطة يمكن ان يفهمها حتى الشخص غير المتخصص بكل سهولة.

وبالرجوع الى بيانات المثال السابق، يمكن تمثيله على شكل أعمدة بيانية كما في الشكل الآتي.



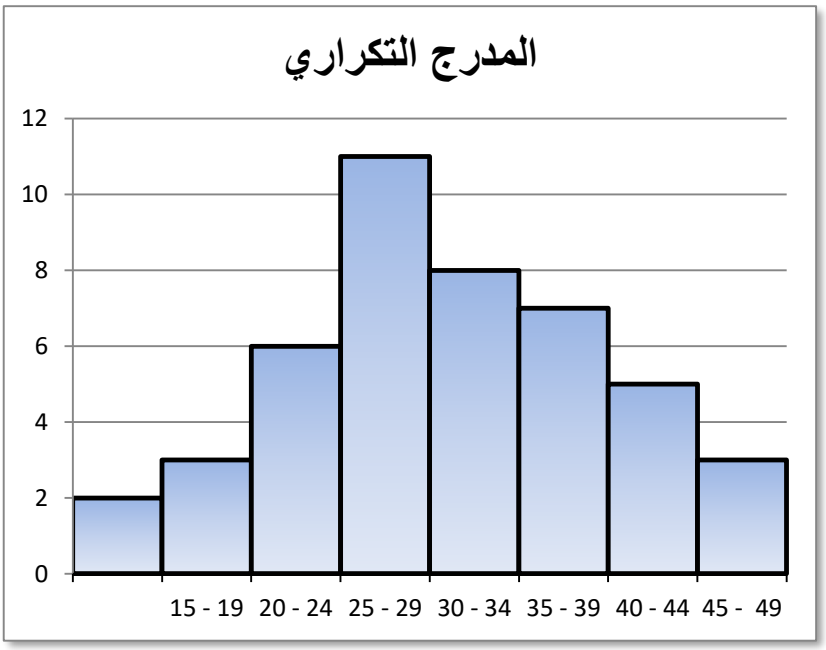
المدرج التكراري: مجموعة من المستطيلات أو الأعمدة المتلاصقة التي يمثل كل عمود عدد التكرارات التي تنتمي لتلك الفئة وأن المجموع الكلي لهذه الأمثلة يمثل الظاهرة أو العينة. وتحدد الفئات على المحور الأفقي (x) في حين تحدد التكرارات على المحور العمودي (y)، ويمكن للمدرجات التكرارية الإجابة على الأسئلة الخاصة بالبيانات المعروضة، مثل ما هو توزيع القيم الرقمية وتكرار حدوثه في مجموعة البيانات؟ هل توجد قيم شاذة؟ فهو يمكننا من فهم البيانات وتوزيعها وبالتالي يمكننا من تحليل البيانات والوصول إلى قرارات مهمة.

مثال: قام معلم بتطبيق اختبار تحصيلي على مجموعة من تلاميذ الصف الخامس وكانت درجاتهم كما في الجدول الآتي:

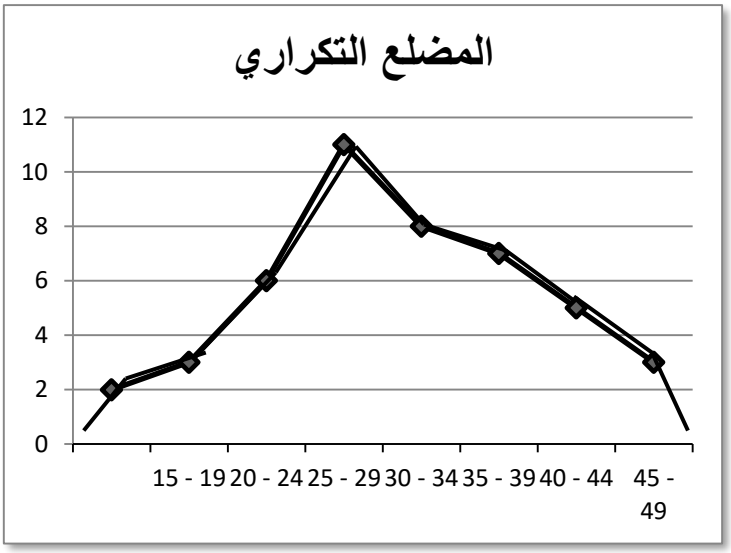
الفئات	التكرارات (f)
14 -10	2
19 -15	3
24 -20	6
29 -25	11
34 -30	8
39 -35	7
44 -40	5
49 -45	3
المجموع	45

ولكي نصمم مدرج تكراري يوضح البيانات الموضحة في الجدول، فإننا نتبع الخطوات الآتية:

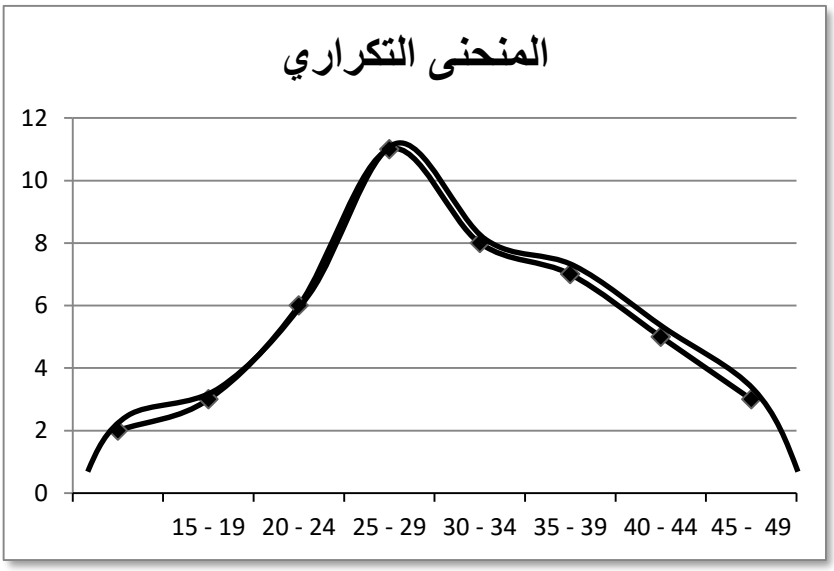
- نرسم المحورين الأفقي (X) والعمودي (y).
- نقسم المحور الأفقي (X) الى أجزاء متساوية ليمثل كل جزء منها طول الفئة ثم نسجل قيمة الحد الأدنى الحقيقي للفئة الصغرى وهي في هذه الحالة (9.5) كأول نقطة من نقاط التقسيم والاستمرار تصاعديا من اليسار الى اليمين حتى الوصول الى الحد الأعلى للفئة الأخيرة ومقدارها في هذه الحالة (49.5) بإضافة (0.5) الى الحد الأعلى للفئة الأخيرة.
- نقسم المحور العمودي (y) الى أقسام متساوية آخذين بالحسبان الحد الأعلى للتكرار وهو في هذه الحالة (12) والحد الأدنى للتكرار وهو في هذه الحالة (2) ويفضل في مثالنا هذا ان تكون المسافة بين كل قسم والذي يليه مقدارها (5).
- نقوم بإنشاء مستطيلات قاعدة كل منها بطول إحدى الفئات وارتفاعه يساوي عدد التكرارات المقابلة لتلك الفئة مع ملاحظة الحدين الأدنى والأعلى الحقيقيين لكل فئة. وكما موضح في الشكل الآتي.



المضلع التكراري: لرسم المضلع التكراري الخاص بتوزيع تكراري معين نستخدم مراكز الفئات لتحديد النقاط الخاصة بتكرار كل فئة ثم يتم ربط هذه النقاط بعد تحديدها جميعا بخطوط مستقيمة فيتكون لدينا شكل عبارة عن خطوط مستقيمة بدلا من المستطيلات كما هو الحال في المدرج التكراري فلكي نرسم مضلعا تكراريا للبيانات المعروضة في الجدول السابق نستخرج مراكز الفئات ثم نحدد موقعها على المحور الأفقي (X) ثم نبدأ بتعيين النقطة الأولى، وهي التي تتكون من التقاء العمود المقام على المحور الأفقي عند مركز الفئة الصغرى وهي (١٠-١٤) ومقداره (١٢) مع العمود المقام على المحور العمودي (y)عند تكرار هذه الفئة وهو (٥) وهكذا نقوم بنفس العملية بالنسبة لجميع الفئات ثم نصل بين كل نقطة والنقطة التالية لها بخط مستقيم فيتكون لدينا المضلع التكراري الموضح في الشكل الآتي, ويمكن غلق المضلع بإضافة فئتين للمحور الأفقي إحداهما سابقة للفئة الأولى والأخرى تالية للفئة الأخيرة واعتبار تكرار كل منهما صفر ثم نربط النقطة الممثلة لمركز الفئة الأخيرة بمركز الفئة التالية لها.



<u>المنحنى التكراري</u>: إن فكرة المنحنى التكراري نفس فكرة المضلع التكراري ونعين النقاط على المحورين، مركز الفئة مع تكرارها أو من المدرج التكراري بتوصيل نقط منتصفات القواعد العليا للمستطيلات المكونة للمدرج التكراري ورسم (منحنى) يمر بأكبر عدد ممكن من النقاط وبتوازن خلال باقي النقاط، ويمكن تحويل الشكل السابق الى منحنى تكراري، وكما هو مبين بالشكل الآتي:



الدائرة البيانية: او ما يسمى بالمخططات الدائرية ويمكن استخدام هذا النوع عندما تكون لدينا مجموعة معينة من البيانات ذات أنواع او مستويات متعددة يمثل كل نوع منها (جزءا فرعيا) من (المجموع العام) للبيانات فعندما تكون لدينا مثل هذه البيانات يمكن عرضها بشكل دائرة مقسمة الى أجزاء متعددة يمثل كل جزء من الدائرة نوعا معينا من البيانات حيث تتناسب مساحة ذلك الجزء من الدائرة مع حجم البيانات الذي تمثله كما يمكن استخدام مختلف الأساليب لجعل هذه الأشكال الدائرية جذابة تجلب انتباه الأشخاص الملاحظين ومن أمثلة هذه الأساليب تمييز أجزاء الدائرة بواسطة ألوان خاصة لكل جزء منها او بواسطة تخطيطها او تظليلها.

أما كيفية عرض البيانات بواسطة هذه الأشكال الدائرية فتتم بواسطة معرفة النسبة المئوية لكل جزء الى المجموع الكلي ومن ثم يتم تمثيل (1%) من البيانات بقطاع تبلغ زاويته المركزية (3.6) درجة وذلك على اعتبار ان قيمة الزاوية المركزية للدائرة هي (360) درجة.

وكمثال: على ذلك لنفرض ان عدد التلاميذ المسجلين في الصفوف الابتدائية في إحدى مديريات التربية لأحد الأعوام الدراسية، كما موضح في الجدول الآتي.

ت	الصف	عدد التلاميذ	النسبة المئوية	مقدار الزاوية
١	الأول	4750	24%	86.4
٢	الثاني	4250	22%	79.2
٣	الثالث	3750	19%	68.4
٤	الرابع	2750	14%	50.4
٥	الخامس	2250	11%	39.6
٦	السادس	2000	10%	36
	المجموع	19750	100%	360

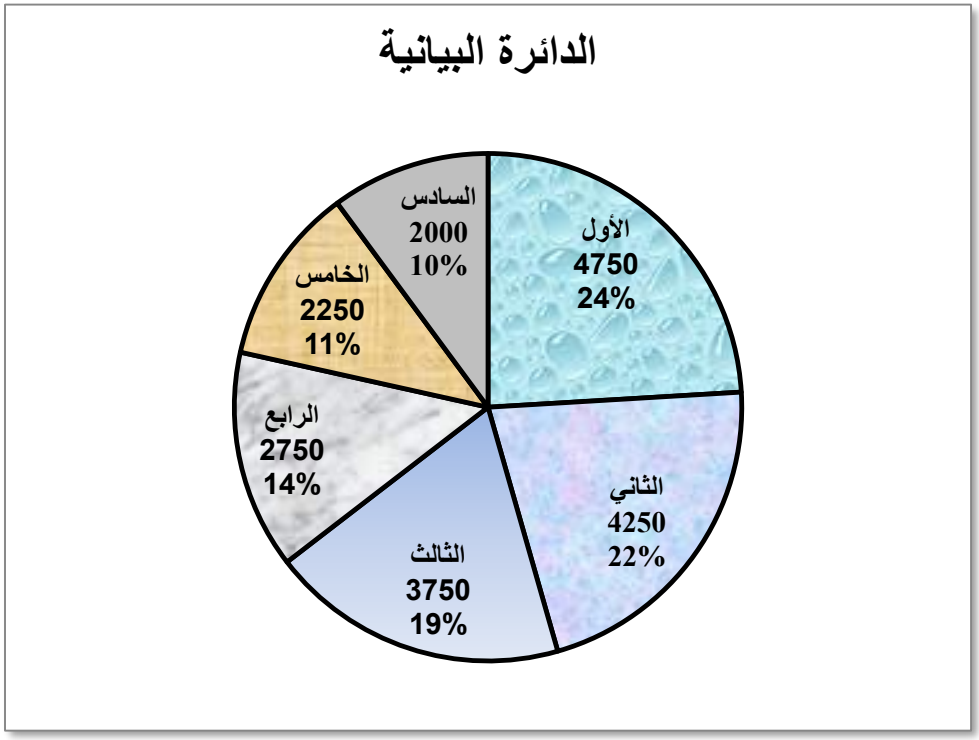
ولعرض هذه البيانات على شكل دائرة بيانية، نحسب النسبة المؤية لتلاميذ كل سنة الى مجموع التلاميذ الكلي لكل السنوات وذلك بقسمة عددهم على المجموع الكلي مضروبا في (100)، وكما يأتي:

الصف الأول = $\frac{4750}{19750}$ x 100= %24

وهكذا نقوم بنفس العملية بالنسبة لتلاميذ الصفوف الأخرى وبذلك يكون مجموع النسب (100%)، ثم نقوم بضرب كل نسبة من هذه النسب في (360):

الصف الأول = 86.4 = %24 x 360

فتتكون لدينا قيم الزوايا لكل جزء من أجزاء الدوائر الستة. وكما موضح في الشكل الآتي.



الفصل الثالث

- مقياس النزعة المركزية.
- الوسط الحسابي (المتوسط).
- الوسط المرجح (الموزون).
- الوسيط.
- المنوال.

مقياس النزعة المركزية.

الاعتماد على الأسلوب البياني وحدة لا يكفي في إعطاء صورة واضحة عن الظاهرة المدروسة، كما وان تحليل ودراسة الظواهر لتحديد الخصائص والاتجاهات والعلاقات، يعتمد في دقته على دقة التمثيل البياني نفسه وبذلك ربما تختلف الخصائص من رسم إلى آخر لنفس الظاهرة، وعليه فإنه من الأفضل اللجوء إلى طرق القياس الكمي، حيث يستخدم الباحث الطريقة الرياضية في القياس. إذ يحتاج الباحث والمعلم في الكثير من النواحي التطبيقية إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تنزع إليها القيم، وهذه المؤشرات تسمى بمقاييس النزعة المركزية او مقاييس الموضع أو المتوسطات.

ومقاييس النزعة المركزية هي مقاييس عددية تستخدم لقياس وتحديد القيم التي تتركز بقية القيم حولها او موضع تركز أو تجمع البيانات. إذ أن بيانات أي ظاهر تنزع في الغالب إلى التركز والتجمع حول قيم معينة. ومقاييس النزعة المركزية تستخدم لتلخيص البيانات عدديا إذ أنها تعتبر قيم نموذجية أو مثالية للبيانات. كما أن هذه المقاييس تستخدم لوصف مجموعة البيانات وكذلك لمقارنة مجموعات البيانات المختلفة. ومن أهم هذه المقاييس نذكر: الوسط أو المتوسط الحسابي، الوسط الموزون أو المرجح، الوسيط، والمنوال.

١- الوسط (المتوسط) الحسابي (Mean).

الوسط أو المتوسط الحسابي من أهم مقاييس النزعة المركزية، وأكثرها استخداما في الإحصاء والنواحي التطبيقية العملية، إذ يستخدم عادة في الكثير من المقارنات بين الظواهر المختلفة. ويعرف الوسط الحسابي بشكل عام على أنه (مجموع القيم مقسوما على عددها). ويرمز لها بالرمز ($\bar{x}$).

فإذا كان لدينا (n) من القيم ،: $x_1, x_2, \ldots, x_n$. فإن الوسط الحسابي لهذه القيم، يحسب بالمعادلة الآتية:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum x}{n}$$

ويمكن حساب الوسط الحسابي للبيانات غير المبوبة، والمبوبة، وكما يأتي:

<u>اولا: حساب الوسط الحسابي للبيانات غير المبوبة.</u> مثلاً، فيما يلي درجات (8) تلاميذ في اختبار صفي (34- 32 - 42 - 37 - 35 - 40 - 36 - 40)، والمطلوب إيجاد الوسط

الحسابي.

الحل: لإيجاد الوسط الحسابي للدرجات نطبق معادلة الوسط الحسابي وكما يأتي:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$= \frac{34+32+42+37+35+40+36+40}{8} = \frac{296}{8} = 37$$

أي أن الوسط الحسابي لدرجة التلميذ في هذا الاختبار يساوي (37= $\bar{x}$) درجة.

ثانيا: الوسط الحسابي للبيانات المبوبة: في حالة البيانات المبوبة توضع القيم الأصلية في شكل فئات، ولا يمكن معرفتها من جدول التوزيع التكرار، ولذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز هذه الفئة، ومن ثم يؤخذ في الحسبان أن (مركز الفئة) هو القيمة التقديرية لكل قيمة تقع في هذه الفئة. فإذا كانت (k) هي عدد الفئات، وكانت (x_1 ,x_2 ,.... ،x_k) هي مراكز هذه الفئات، و (f_1 ,f_2 ,.... ،f_k) هي التكرارات، فإن الوسط الحسابي يحسب بالمعادلة الآتية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_k f_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} \qquad (٢-٣)$$

مثال: الجدول التالي يعرض توزيع (40) تلميذ حسب أوزانهم. والمطلوب إيجاد الوسط الحسابي.

فئات الوزن	47-49	44-46	41-43	38-40	35-37	32-34
عدد التلاميذ	1	5	10	13	7	4

الحل، لحساب الوسط الحسابي نستخدم المعادلة السابقة، ويتم إتباع الخطوات الآتية:

١. إيجاد مجموع التكرارات ($\sum f$). ٢. حساب مراكز الفئات (x= 3). ٣. ضرب مركز الفئة في التكرار المناظر له ($x.f$)، وحساب المجموع ($\sum x.\ f$). ٤.حساب الوسط الحسابي بتطبيق المعادلة:

فئات الوزن (C)	التكرارات f	مراكز الفئات x	$(x.f)$
32-34	4	32+34/2=33	132= 33 x 4
35-37	7	36	252=36 x 7
38-40	13	39	507=39 x13
41-43	10	42	420=42 x10
44-46	5	45	225=45x5
47-49	1	48	48=48x1
المجموع	40		1584

إذا الوسط الحسابي لوزن التلميذ هو:

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i f_i}{\sum_{i=1}^{6} f_i} = \frac{1584}{40} = 39.6\ k.g$$

أي أن متوسط وزن التلميذ يساوي (39.6 k. g)

خواص الوسط الحسابي:

مميزات الوسط: إن المتوسط يعتبر من أفضل مقاييس النزعة المركزية ومن أكثرها شيوعا وذلك لما يتمتع به من صفات جيدة. ومن مميزات المتوسط نذكر ما يلي:

1. المتوسط سهل التعريف والحساب ويخضع للعمليات الجبرية بسهولة.
2. المتوسط وحيد لمجموعة البيانات الواحدة. وهو نقطة اتزان المشاهدات.
3. المجموع الجبري لانحرافات القيم عن وسطها يساوي صفرا.
4. مجموع مربعات الانحرافات عن الوسط الحسابي يكون أصغر من مجموع مربعات الانحرافات عن أي قيمة أخرى.
5. يأخذ المتوسط في الاعتبار جميع البيانات (جميع القيم والمشاهدات).
6. اقل مقاييس النزعة المركزية تأثرا بالتقلبات العينية.

عيوب الوسط: لا يصلح المتوسط الحسابي في هذه الحالات:

1. عدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.
2. يتأثر بالقيم المتطرفة والقيم الشاذة لذا لا يصلح للتوزيعات الملتوية بمعنى أن الوسط الحسابي للقيم (٢٠، ٣٠، ٤٠ ،٥٠، ١٢٠) هو ٢٦٠ ÷ ٥ = ٥٢ أكبر من غالبية القيم الموجودة بسبب القيمة (١٢٠) الشاذة عن باقي القيم الأخرى المتقاربة فيما بينها وهذا يقودنا لعدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.
3. لا يمكن استعماله مع البيانات المبوبة في فئات وكان التوزيع مفتوحا في أحد طرفيه.
4. لا يمكن إيجاده من الرسم البياني كالوسيط والمنوال.
5. لا يمكن إيجاده من التوزيع التكراري المفتوح إلا في حال تقدير مراكز الفئات المفتوحة.

٢ – الوسط المرجح (الموزون) Weighted Mean.

في بعض الأحيان تكون المشاهدات ($X_1, X_2, \ldots, X_n$) مقرونة بالأوزان ($W_1, W_2, \ldots, W_n$) على التوالي. وفي هذه الحالة نعرف الوسط المرجح كما في المعادلة الآتية:

$$\bar{X}_w = \frac{\sum w_i x_i}{\sum w_i} = \frac{x_1 w_1 + x_2 w_2 + \cdots + x_n w_n}{w_1 + w_2 + \cdots + w_n}$$

Ɣŏ̈oʷ̈! ɬɳkl̨dž!ƒ ƧƓlhŮƒ LJdž YƒĂl̨ƒ ɬ ŭλ! yɳƧdž!ƒ ƱƧĂ!ƒ ƥɳĂƒ ýƓ؞

انواع الابقار	نسبة الدهن	كمية الانتاج
فريزان	4	65
هورشتاين	3	70
الشرابي	2	40

الحل:

x	w	wx
65	4	260
40	2	80
70	3	210
	∑w= 9	∑ wx = 550

$$\bar{x}_w = \frac{\sum x\,w}{\sum w} = = \frac{550}{9} = 61.11$$

الوسط المرجح = (61.11)

٣- الوسيط (Median).

هو أحد مقاييس النزعة المركزية، والذي يأخذ في الحسبان رتب القيم، أي، يعتمد الوسيط على ترتيب القيم والمفردات وليس على قيم المفردات ذاتها لذلك يجب ان ترتب الدرجات تصاعديا او تنازلي. ويفضل استعماله في البيانات التي توجد فيها تطرف او في البيانات التي تكون تكراراتها مفتوحة. ويعرف الوسيط لمجموعة من البيانات على أنه تلك النقطة او القيمة التي تتوسط مجموعة من البيانات المرتبة (تصاعدياً او تنازلياً) ويقل عنها نصف عدد القيم، ويزيد عنها النصف الآخر، أي أن (50%) من القيم تساويه او أقل منه، و(50%) من القيم تساويه او أعلى منه. ويرمز للوسيط بالرمز (Med). وفيما يأتي كيفية حساب الوسيط في حالة البيانات غير مبوبة، والبيانات المبوبة.

أولا: الوسيط للبيانات غير المبوبة: لبيان كيف يمكن حساب الوسيط للبيانات غير المبوبة، نتبع الخطوات الآتية:

١- ترتيب القيم تصاعديا.

٢- تحديد رتبة الوسيط:

أ- إذا كان عدد القيم (n) فردي فإن رتبة الوسيط هي: ($R_{me}=n+1/2$).

مثال: البيانات الآتية تتكون من (n) تسع قيم مرتبة تصاعديا.

القيم	3	4	6	7	8	11	12	14	17
الرتب	1	2	3	4	5	6	7	8	9

ولتحديد رتبة الوسيط نعتمد المعادلة الخاصة بعدد القيم الفردي.

($R_{me}=9+1/2$ +1=10/2=5), ونجد ان القيمة (9) تقع في الرتبة الخامسة (5).

ب- وإذا كان عدد القيم (n) زوجي، فإن الوسيط يقع بين القيمة بالرتبة ($n/2$)، والقيمة بالرتبـــــة [1+ ($n/2$)]، ومن ثم يحسب رتبة الوسيط:

مثال :البيانات الآتية تتكون من (n) عشرة قيم مرتبة تصاعديا .

القيم	3	4	6	7	9	11	12	14	15	17
الرتب	1	2	3	4	5	6	7	8	9	10

ولتحديد رتبة الوسيط نعتمد المعادلة الخاصة بعدد القيم الزوجي. يكون ترتيب الوسيط بين الرتبتين (5 = 2 /10)، والرتبة [6 =1+ (10/2)]. أي ان رتبتي الوسيط (5 – 6) نجد القيمتين (9 – 11) في هاتين الرتبتين، أي ان قيمة الوسيط يقع في الرتبة (5.5)، ويساوي ($Me = 9 + 11 = 20/2= 10$).

ثاني: الوسيط للبيانات المبوبة: يتم حساب قيمة الوسيط من بيانات مبوبة في جدول توزيع تكراري، وفقا للمعادلة:

$$me = A + \frac{\frac{n}{2} - F1}{F2} \times L$$

ثم يتم إتباع الخطوات الآتية.

١- تكوين الجدول التكراري المتجمع الصاعد.

٢- تحديد رتبة الوسيط: $\frac{\sum f}{2} = \frac{n}{2}$

٣- تحديد القيم: A = الحد الأدنى الحقيقي للفئة الوسيطية. L = طول الفئات.

F1 = التكرار المتجمع الصاعد للفئة السابقة للفئة الوسيطية.

F2 = تكرار الفئة الوسيطية.

٤- نوجد الوسيط بالعلاقة السابقة:

مثال: احسب قيمة الوسيط للبيانات في الجدول الآتي:

الفئات (C)	التكرارات (f)	التكرار المتجمع الصاعد (f↓)
12 - 10	2	2
15 - 13	3	F_1 5
18 - 16	F2 7	12
21 - 19	4	16
24 - 22	2	18
المجموع ∑	18	

الحل: ١- نحدد رتبة الوسيط (R_{me}= 18/2 = 9)

٢- ثم نحدد الفئة الوسيطية وذلك بان نحدد التكرار المتجمع الصاعد الذي يتضمن رتبة الوسيط (9)، وهو التكرار المتجمع الصاعد (12) والذي يقابل الفئة (16 – 18).

٣- نحدد القيم: رتبة الوسيط ((n/2 = 18/2= 9) ..... (A=16-0.5 = 15.5) ..... (L = 3) ..... (F1=5) ..... (F2= 7). ثم نطبق القانون:

$$me = A + \frac{\frac{n}{2} - F1}{F2} \times L = me = 15.5 + \frac{9-5}{7} \times 3 = 15.5 + \frac{4}{7} \times 3$$

$$me = 15.5 + 0.8 \times 3 \ = 15.5 + 2.4 = \underline{me = 17.9}$$

ملاحظة :يجب ان تكون قيمة الوسيط ضمن حدود الفئة الوسيطية.

بعض مميزات وعيوب الوسيط:

مميزات الوسيط: يعد الوسيط من مقاييس النزعة المركزية الشائعة وذلك لما يتمتع به من بعض الصفات الجيدة. ومن مميزات الوسيط نذكر ما يلي:

١. لا يتأثرا من المتوسط بالقيم الشاذة أو المتطرفة، لذلك يستخدم بدل من المتوسط الحسابي في مثل هذه الحالات.

٢. لا تتأثر قيمة الوسيط كثيراً عند إعادة التوزيع التكراري.

٣. يمكن استخدامه في حالة الجداول ذات الفئات المفتوحة لأنه لا يعتمد على مراكز الفئات.

عيوب الوسيط: وكما للوسيط مميزات جيدة إلا أن له بعض العيوب نذكر منها ما يأتي:

١. لا يأخذ الوسيط في الاعتبار جميع البيانات إذا أنه يعتمد فقط على القيم التي في المنتصف وعلى ترتيب البيانات بغض النظر عن قيمها.

٢. تختلف قيم الوسيط من عينة إلى أخرى لنفس المجتمع بعكس المتوسط.

٣. اقل دقة من المتوسط الحسابي في وصف طبيعة البيانات. لأنها تتناول فقط برتبة البيانات وليس قيمها.

٥- المنوال Mode.

يعد المنوال من أسهل مقاييس النزعة المركزية، ويمكن تحديد المنوال دون الحاجة الى إجراءات إحصائية معقدة، ويعرف المنوال بأنه القيمة او المشاهدة الأكثر شيوعا أو تكرارا، ويكثر استخدامه في حالة البيانات (الوصفية)، لمعرفة النمط (المستوى) الشائع، ويرمز بالرمز (Mo) ويمكن حسابه للبيانات المبوبة وغير المبوبة كما يلي:

أولا: حساب المنوال في حالة البيانات غير المبوبة: في الجدول الآتي حالات مختلفة للمنوال.

20	18	16	15	11	9	8	6	5	x	1
-	-	-	-	-	-	-	-	-	Mo	
16	13	11	9	9	9	8	7	5	x	2
-	-	-	Mo			-	-	-	Mo	
19	16	14	12	12	10	10	9	7	x	3
-	-	-	(10 +12) /2 = 11				-	-	Mo	
20	19	17	17	16	15	14	14	12	x	4
-	-	Mo 2		-	-	Mo 1		-	Mo	
19	16	16	15	10	10	8	5	5	x	5
-	-	-	-	-	-	-	-	-	Mo	

نلاحظ من الجدول ان البيانات برقم:

(١) لا يوجد منوال، لأنه لا توجد قيمة أكثر تكرارا.

٢) القيمة المنوالية (9) الأكثر تكرارا.

(٣) يوجد منوالين متجاورين (10) و(12)، في هذه الحالة تجمع قيمتي المنوالين وتقسم على (11 = 2 / (12+ 10)) لتصبح القيمة المنوالية (11).

(٤) توجد قيمتين منواليتين غير متجاورتين المنوال الأول (14)، والمنوال الثاني (17). ويسمى التوزيع ذو المنوالين.

(٥) يوجد أكثر من قيمة بنفس التكرار، إذا لا يوجد منوال.

ثانيا: حساب المنوال في حالة البيانات المبوبة. يتم حساب المنوال من بيانات مبوبة في جدول توزيع تكراري، وفقا للمعادلة:

$$mo = A + \frac{D1}{D1 + D2} \times L$$

ثم يتم إتباع الخطوات الآتية.

١- تحديد الفئة المنوالية، وهي الفئة الأكثر تكرارا.

٢- تحديد القيم: A = الحد الأدنى للفئة المنوالية. L = طول الفئات.

D1 = الفرق بين تكرار الفئة المنوالية وتكرار الفئة قبل المنوالية.

D2 = الفرق بين تكرار الفئة المنوالية وتكرار الفئة بعد المنوالية.

٤- نحسب المنوال وفقا للعلاقة السابقة:

مثال: احسب قيمة المنوال للبيانات في الجدول الآتي:

الفئات (C)	التكرارات (f)	
24 - 20	4	
29- 25	7	D1= 9-7= 2
34 - 30	9	
39 - 35	5	D2= 9-5= 4
44 - 40	4	
49 - 45	3	

الحل: ١- نحدد الفئة المنوالية وهي (٣٤ - ٣٠) الأكثر تكرارا (٩).

٢- نحدد القيم: (A = 30- 0.5= 29.5)، (L=5)، (D1= 9-7= 2)، (D2= 9-5= 4).

٣- نطبق معادلة المنوال:

$$mo = A + \frac{D1}{D1+D2} \times L = \; mo = 29.5 + \frac{2}{2+4} \times 5 = 29.5 + \frac{2}{6} \times 5$$

$$= 29.5 + 0.33 \times 5 = 29.5 + 1.65 = \underline{mo = 31.15}$$

نلاحظ ان المنوال بلغ (31.15) وهو ضمن الحدود الحقيقية للفئة المنوالية.(30 - 34)

مميزات وعيوب المنوال:

مزايا المنوال: يعد المنوال من مقاييس النزعة المركزية الشائعة ومن مميزاته:

1. انه مقياس سهل الفهم والحساب.
2. يكن تقدير المنوال عن طريق ملاحظة التكرار الأكثر.
3. يمكن حساب المنوال للبيانات الكمية والوصفية (النوعية).
4. ان المنوال لا يتأثر إطلاقا بالقيم الشاذة والمتطرفة.
5. يمكن إيجاد المنوال في حالة التوزيعات التكرارية المفتوحة من طرف واحد او طرفين.
6. يمكن تعيين المنوال هندسيا.

ب- عيوب المنوال: بالرغم من أن المنوال يعتبر من مقاييس النزعة المركزية الشائعة إلا أن له

بعض العيوب:

1. يتأثر المنوال وعلى نحو كبير بأخطاء المعاينة.
2. لا تستند عملية إيجاده إلى كافة البيانات المتاحة، حيث انه بمجرد ملاحظة أكبر تكرار يتم معرفة المنوال او فئته وعندئذ تمهل كافة القيم الأخرى أو الفئات الأخرى.
3. لا يخضع للعمليات الحسابية الجبرية.
4. قد لا يوجد منوال لمجموعة من البيانات أو قد يكون هناك أكثر من منوال.

مقارنة بين الوسط الحسابي والوسيط والمنوال.

1. يعد الوسط الحسابي أحسن مقاييس النزعة المركزية ملائمة في حالة البيانات الفاصلية والنسبية على حين يفضل استعمال الوسيط في حالة البيانات الرتبية والمنوال للبيانات الاسمية.
2. ان حساب الوسط الحسابي يتضمن كل درجة من درجات المجموعة، ويتأثر الوسط الحسابي إذا طرا أي تغير لأي درجة، وعلى العكس من ذلك الوسيط لا يعتمد ولا يتأثر بتغير الدرجات وإنما يتأثر بتغير عددها.
3. لا يتأثر المنوال بالدرجات المتطرفة كما هو الحال في الوسط الحسابي ولا يتأثر بالدرجات الوسطى كما هو الحال في الوسيط ولذلك فهو من هذه الناحية أكثر استقرارا من الوسط الحسابي والوسيط.
4. تتساوى قيم الوسط الحسابي والوسيط والمنوال عندما يكون منحني التوزيع متماثلا وهو المنحني الذي يتطابق نصفاه.

الفصل الرابع

- مقياس التشتت.
- الانحراف المعياري.
- التباين.
- الدرجة المعيارية

مقاييس التشتت Dispersion Measurements:

ان مقاييس النزعة المركزية تساعد الى حد بعيد في التعرف على مدى تجمع او تمركز مجموعة من الدرجات حول مقياس عددي معين. فمقياس النزعة المركزية تعطي قيمة واحدة، لتكون ممثلة الى حد ما لجميع القيم الموجودة في مجموعة البيانات (الدرجات)، وبذلك تعبر عن المستوى العام للظاهرة محل البحث والدراسة. بغض النظر عن الفروق الموجودة بين تلك الدرجات.

وهكذا نلاحظ ان استعمال مقاييس التمركز بمفردها غير كافية لوصف التوزيع لمجموعة من البيانات وصفاً كاملاً ومعرفة خصائصه بشيء من الدقة. لان هذه المقاييس لا تصف مدى تقارب وتباعد قيم البيانات عن بعضها البعض، أو مدى تباعد أو تقارب هذه القيم عن مركزها. ولذلك تبرز الحاجة لوصف التوزيع من حيث التعرف على مدى انتشار الدرجات فيه واختلافها الواحدة عن الأخرى. وكذلك الحال عند المقارنة، فقد يكون مقياس التمركز للمجموعتين متساوي، وربما يوجد اختلاف كبير في درجة تجانس قيم البيانات بين المجموعتين من حيث مدى تقاربها وتباعدها. إذا، نحتاج الى مقاييس التشتت تقيس مدى التجانس لمجموعة الدرجات لتكون لدينا صورة أوضح لها. وأهم مقاييس التشتت او الاختلاف هي: المدى، والانحراف المعياري، والتباين، والدرجة المعيارية.

المدى Range:

المدى هو ابسط المقاييس الإحصائية في التعرف على مدى تشتت مجموعة قيم البيانات او درجات التوزيع، ويعرف المدى بأنه مقدار الفرق الموجود بين أكبر درجة وأصغر درجة في البيانات او التوزيع. وهناك نوعين من المدى: المدى المطلق، والمدى الحقيقي.

١- **المدى المطلق**: هو الفرق الموجود بين أكبر درجة وأصغر درجة في المجموعة. وتحسب حسب المعادلة ($R = X_L - X_s$).

٢- **المدى الفعلي (الحقيقي)**: هو الفرق الموجود بين الحد الحقيقي لأكبر درجة والحد الحقيقي لأصغر درجة في المجموعة، وذلك بإضافة (0.5) درجة لأكبر درجة وطرح (0.5) درجة لأصغر درجة، أي إضافة درجة لناتج الفرق. وتحسب حسب المعادلة ($R = X_L - X_s + 1$).

مثال: لحساب المدى المطلق والحقيقي للبيانات الآتية والمرتبة تصاعديا:

36	35	32	31	30	28	26	24	22	21

الحل: المدى المطلق يساوي (15 =21 – 36)، أما المدى الحقيقي يساوي (16 = 1+ 21 – 36).

مميزات وعيوب المدى: ومما تقدم ان المدى يتصف بأنه:

المميزات: المدى يتميز بأنه:

١- سهل الحساب.

٢- يعطي فكرة سريعة عن طبيعة البيانات.

العيوب: ومن عيوب المدى:

١- يعتمد في حسابه على قيمتين فقط (أكبر قيمة وأصغر قيمة) وأنه يهمل بالتالي باقي القيم.

٢- لا يقيس تشتت البيانات عن متوسطها، فهو لا يشير – من قريب أو من بعيد – إلى متوسط البيانات (أو مركزها).

٣- حساس جداً لأي قيمة شاذة أو متطرفة، ولا يأخذ بنظر الاعتبار القيم الأخرى في التوزيع. مما يؤدي إلى إعطاء صورة غير صحيحة عن البيانات الإحصائية التي نقوم بوصفها وتفسيرها.

التباين The Variance.

يعد التباين من مقاييس التشتت التي تعتمد على كل قيمة (درجة) من درجات التوزيع، ومدى انحرافها عن الوسط الحسابي. وهذا المقياس مهم لأنه من ناحية يأخذ جميع القيم في الحسبان، ومن ناحية أخرى لأنه يقيس التشتت عن الوسط الحسابي للقيم، هذا بالإضافة إلى أنه تسهل معالجته رياضياً، وأنه يدخل في تكوين عدد من المقاييس والاختبارات الإحصائية المهمة.

والفكرة الأساسية للتباين هي حساب انحرافات جميع القيم (الدرجات) عن وسطها الحسابي (أي حساب الفرق بين كل قيمة والوسط الحسابي)، وسوف نلاحظ أن بعض الفروق أو الانحرافات موجبة (القيم أكبر من الوسط)، وبعض الفروق (أو الانحرافات) سالبة (القيم أصغر من الوسط). ويكون مجموع هذه الانحرافات مساوياً للصفر في جميع الحالات. ويكون الحل هنا إما إهمال او حذف الإشارات السالبة أو تربيع هذه الانحرافات. وحذف الإشارات السالبة عن القيم غير صحيح، وليس له مبرر رياضي، ويصعب التعامل معها، فيكون الحل هو تربيع هذه الانحرافات، ثم نحسب متوسط الانحرافات المربعة فنحصل على التباين. أي أن التباين يعرف بأنه متوسط مربعات انحرافات القيم عن وسطها الحسابي. ويحسب التباين (S^2) بطريقتين: الانحرافات، والتربيعات حسب المعادلتين:

$$S^2=\frac{x^2-(\sum x)^2/n}{n-1}$$ (بطريقة التربيعات) $$S^2=\frac{(\bar{x}-x)^2}{n-1}$$ ..... (بطريقة الانحرافات)

مثال :احسب تباين الدرجات.(13 -17 -16 -14 -15 -12 -15 -10)

الحل: لحساب التباين بطريقة الانحرافات نتبع الخطوات الآتية:

١- نستخرج الوسط الحسابي لهذه الدرجات

$$\bar{x} = \frac{13+17+16+14+15+12+15+10}{8} = \frac{112}{8} = \underline{14}$$

٢- نستخرج انحرافات القيم عن الوسط الحسابي (أي الفروق بينها وبين الوسط الحسابي), وذلك بطرح الوسط الحسابي من كل قيمة ($\bar{x}$ - x). ثم نربع هذه الانحرافات.

٣- ثم نجمع مربعات الانحرافات وتقسم على عددها كما في الجدول الآتي:

x	$(x - \bar{x})$	$(x - \bar{x})^2$
13	-1 =14 - 13	1
17	3 =14 - 17	9
16	2 =14 - 16	4
14	0 =14 - 14	0
15	1 =14 - 15	1
12	-2 =14 - 12	4
15	1 =14 - 15	1
10	-4 =14 - 10	16
$\sum$	0	36

٤- ثم نحسب التباين $S^2 = \frac{36}{8-1} = \frac{36}{7} = \underline{5.14}$

ولحساب التباين بطريقة المربعات نتبع الخطوات الآتية:

١- نحسب مجموع القيم (الدرجات).

٢- نربع القيم، ثم نجمع مربعات الدرجات. كما في الجدول:

n	x	x^2
1	13	169
2	17	289
3	16	256
4	14	196
5	15	225
6	12	144
7	15	225
8	10	100
$\sum$	112	1604

٣- نحسب التباين وفق المعادلة:

$$S^2 = = \frac{x^2-(\sum x)^2/n}{n-1} = = \frac{1604-(112)^2/8}{8-1} = \frac{1604-12544/8}{7} = \frac{1604-1568}{7} = \frac{36}{7}$$

$S^2 = \underline{\underline{5.14}}$

الانحراف المعياري.Standard Deviation

يعرف الانحراف المعياري بأنه الجذر ألتربيعي الموجب للتباين. والانحراف المعياري عادة ما يستخدم بدلاً من التباين كأهم مقياس للتشتت. وذلك لان قيمته تكون أصغر – جذر – التباين. ويستخرج الانحراف المعياري (S) بنفس الطرق والمعادلات التي تم استعراضها لاستخراج التباين (S^2). ثم يؤخذ الجذر ألتربيعي للناتج. وذلك حسب المعادلات:

(بطريقة التربيعات) $S= \sqrt{\frac{x^2-(\sum x)^2/n}{n-1}}$ ..... (بطريقة الانحرافات) $S= \sqrt{\frac{(\bar{x}-x)^2}{n-1}}$

بعض خصائص التباين والانحراف المعياري: يشترك مقياسي التباين والانحراف بما يأتي:

١ – يمكن بواسطتهما التعرف على مدى التجانس وتشتت الدرجات في التوزيعات المختلفة، والمقارنة فيما بين تلك التوزيعات او البيانات.

٢– قيمة المقياسين دائماً موجبة أو أكبر من أو تساوي صفر. فأقل قيمة تساوي الصفر (وذلك عندما تكون جميع القيم متساوية، وفي هذه الحالة لا توجد فروق أو انحرافات بينها وبين الوسط الحسابي وبالتالي لا يوجد أي تشتت بين القيم، وبالتالي فإن قيمة التباين او الانحراف المعياري في حالة تساوي جميع القيم تساوي الصفر).

٣ – كلما كان التشتت كبيراً حول الوسط كلما كان التباين والانحراف المعياري كبيراً، والعكس صحيح.

٤ – إذا أضفنا وطرحنا مقداراً ثابتاً من كل القيم في التوزيع فإن قيمة المقياسين لا تتغير (أي لا يتأثر التباين او الانحراف المعياري بالطرح أو الجمع).

٥– للتباين والانحراف المعياري فوائد وخصائص كثيرة في الإحصاء الاستدلالي وفي اختبار الفرضيات.

الدرجات المعيارية Stander Scores.

قد يهدف المعلم ان يحدد موقع وأهمية درجة (ملاحظة) معينة بالنسبة الى مجموعة من الدرجات في نفس التوزيع (على سبيل المثال، امتحان تحصيلي)، او لمقارنة هذه الدرجة مع درجة أخرى في توزيع آخر (امتحان تحصيلي آخر)، ولذلك لتحقيق ذلك يحتاج الى طريقة إحصائية يوحد بها وحدة قياس الدرجة وفي مثل هذه الحالات يمكن تحويل الدرجة (الخام) الأصلية الى ما يسمى بالدرجة المعيارية (الزائية). إذا، الدرجة المعيارية او الزائية هي تعبير كمي عن الدرجة الخام (الملاحظة) لتحديد موقعها النسبي في العينة أو المجتمع. يعني هي عدد الانحرافات المعيارية التي تبتعد عن وسط التوزيع, وتحسب عن طريق قياس الفرق بين الدرجة الخام (x) عن الوسط الحسابي ($\bar{x}$) للتوزيع وقسمته على الانحراف المعياري (S) لنفس التوزيع ويمكن التعبير عن الدرجة المعيارية (Z) رمزيا بما يأتي:

$$Z = \frac{x-\bar{x}}{S}$$

وعلى العموم تكون الدرجة المعيارية (Z) المقابلة للدرجة الخام(X) مساوية للفرق بينها وبين مركز التوزيع مقسوماً على تشتت التوزيع. أي، تأخذ في الحسبان النزعة المركزية (الوسط الحسابي) والتشتت (الانحراف المعياري) لكافة أفراد المجموعة، وتكون الدرجة المعيارية موجبة إذا كانت قيمتها أكبر من الوسط الحسابي وصفراً حال مساواتها بالوسط الحسابي وسالبة إذا كانت قيمتها أقل من الوسط الحسابي.

ويستفاد من الدرجات المعيارية في المقارنة بين الدرجات في نفس التوزيع، وكذلك المقارنة بين الدرجات في توزيعات مختلفة، بشرط معلومية طبيعة هذه التوزيعات من حيث تمركزها (وسطها الحسابي)، ودرجة تشتتها (انحرافها المعياري).

مثال: البيانات في الجدول الآتي لخمسة تلاميذ فقط في اختبارين تحصيلين تم سحبها عشوائيا من صف معين. علما ان الوسط الحسابي والانحراف المعياري يمثل درجات كل تلاميذ الصف، المطلوب حساب الدرجات المعيارية والمقارنة بينها.

تسلسل التلميذ	1	2	3	4	5	.....	الوسط الحسابي ($\ddot{X}$)	الانحراف المعياري (S)
درجات اختبار العلوم (X_1)	50	80	90	60	75	...	60	10
درجات اختبار الرياضيات (X_2)	55	85	90	75	80	...	75	8

الحل: نحسب أولا الدرجات المعيارية المقابلة للدرجات الخام الواردة بالجدول باستخدام معادلة (z):

الدرجة المعيارية للتلميذ بتسلسل(١) في العلوم $Z = \frac{x-\bar{x}}{S} = \frac{50-60}{10} = \frac{-10}{10} = -1$

الدرجة المعيارية للتلميذ بتسلسل(١) في الرياضيات $Z = \frac{x-\bar{x}}{S} = \frac{55-75}{8} = \frac{-20}{10} = -2$

وهكذا نحصل على الدرجة المعيارية للدرجات الخام لبقية التلاميذ، ولتكون الفكرة واضحة ولتسهيل المقارنة نضيف الدرجات المعيارية الى الجدول السابق، ليتكون الجدول الآتي:

تسلسل التلميذ	1	2	3	4	5	الوسط الحسابي $(\bar{x})$	الانحراف المعياري (S)
درجات اختبار العلوم (X_1)	50	80	90	60	75	60	10
الدرجات المعيارية للعلوم(Z_1)	**-1**	**2**	**3**	**0**	**1.5**		
درجات اختبار الرياضيات (X_2)	55	85	90	75	80	75	8
الدرجات المعيارية للرياضيات(Z_2)	**-2.5**	**1.25**	**1.88**	**0**	**0.63**		

نلاحظ من الجدول ان الدرجة الخام (x) لا تعكس واقع صاحبها بصرف النظر عن قيمتها، ويتضح ذلك كما يأتي:

١- درجة تلميذ في التسلسل (١) في العلوم (50)، أصغر من درجته في الرياضيات (55)، إلا ان درجته المعيارية في العلوم (1-) أكبر من درجته المعيارية في الرياضيات (2.5-).

٢- درجة تلميذ في التسلسل (٢) في العلوم (80)، أصغر من درجته في الرياضيات (85)، إلا ان درجته المعيارية في العلوم (2) أكبر من درجته المعيارية في الرياضيات (1.25).

٣- درجة تلميذ في التسلسل (٣) في العلوم (90)، تساوي درجته في الرياضيات (90)، إلا ان درجته المعيارية في العلوم (3) أكبر من درجته المعيارية في الرياضيات (1.88).

٤- درجة تلميذ في التسلسل (٤) في العلوم (60)، أصغر من درجته في الرياضيات (75)، إلا ان درجته المعيارية في العلوم وفي الرياضيات هي ذاتها (0).

٥- درجة تلميذ في التسلسل (٥) في العلوم (75)، أصغر من درجته في الرياضيات (80)، إلا ان درجته المعيارية في العلوم (1.5) أكبر من درجته المعيارية في الرياضيات (0.63).

الفصل الخامس

- **مقاييس الارتباط**.
- **معامل ارتباط بيرسون**.
- **معامل ارتباط سبيرمان**.

مقاييس الارتباط Measures of Correlation

تتطلب العديد من المشكلات والمتغيرات في مجال التربية وعلم النفس دراستها بطرق إحصائية أخرى غير تلك التي تعرفنا عليها سابقا. ومن هذه الطرق الإحصائية مقاييس الارتباط او العلاقة التي تستعمل في دراسة العلاقة والارتباط بين متغيرين، او أكثر بدلا من دراسة متغير واحد.

ويشير مفهوم الارتباط إلى قوة واتجاه العلاقة بين المتغيرات، فقد تكون قوة هذه العلاقة بعدة مستويات (معدومة، ضعيفة، متوسطة، عالية، تامة). كما وان هذه العلاقة قد تكون (طردية). أي، باتجاه (موجب)، او قد تكون علاقة (عكسية). أي، باتجاه (سالب).

وتقاس قيمة الارتباط بمعامل يسمى معامل الارتباط (Coefficient Correlation)، والذي تتراوح قيمته بين (1+, 1−) وكلما قربت قيمة معامل الارتباط من الواحد الصحيح كلما كان الارتباط قوياً، وكلما اقتربت قيمته من الصفر كلما كان الارتباط أضعف. وبذلك يأخذ الارتباط الحالات الآتية:

١- إذا كانت قيمة معامل الارتباط تساوي واحد صحيح موجب (1+)، تعني وجود علاقة خطية (طردية) موجبة تامة بين المتغيرين، بمعنى انه كلما زادت القيم في أحد المتغيرين، تزداد في المتغير الآخر. وتتفاوت قيمة معامل الارتباط الموجب حتى تصل الى الصفر حسب قوة العلاقة.

٢- إذا كانت قيمة معامل الارتباط تساوي واحد صحيح سالب (1−)، تعني وجود علاقة خطية (عكسية) سالبة تامة بين المتغيرين، بمعنى انه كلما زادت القيم في أحد المتغيرين، تقل في المتغير الآخر. وتتفاوت قيمة معامل الارتباط السالب حتى تصل الى الصفر حسب قوة العلاقة.

٣- أما إذا كان قيمة معامل الارتباط مساوي للصفر فان ذلك يدل على انه لاتوجد علاقة بين المتغيرين، وان الارتباط يكون معدوماً.

وتوجد طرق عدة لقياس العلاقة بين متغيرين ويعتمد استعمال هذه الطرق على طبيعة المتغيرين. وفيما يأتي عرض لأهم هذه الطرق:

معامل ارتباط بيرسون Pearson:

وهو إحدى الطرائق الإحصائية المستعملة لقياس درجة العلاقة بين متغيرين مستمرين. فمن الأمور التي يهتم بها الباحثون في مجال التربية وعلم النفس، التعرف على مدى العلاقة الموجودة بين قيم متغيرين مستمرين سواء ان كان كلاهما من النوع الفئوي او النسبي او أحدهما نسبي والآخر فئوي. في هذا الحال يحتاج الباحث الى طريقة لاستخراج معامل الارتباط بين المتغيرين.

وتحسب قيمة واتجاه الارتباط بحسب المعادلة:

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

حيث إن: r = معامل ارتباط بيرسون. n = عدد الدرجات. x, y = قيم المتغيرين.

x^2 , y^2 = مربعات درجات المتغيرين. $\sum$x. y = مجموع حاصل ضرب قيم المتغيرين.

مثال: أراد معلم معرفة العلاقة (الارتباط) بين التحصيل في مادتي اللغة العربية واللغة الانكليزية، ولحقيق ذلك قام باختيار درجات (10) تلاميذ، وكما يأتي:

ت	1	2	3	4	5	6	7	8	9	10
درجات اللغة العربية	7	8	5	7	9	10	5	7	9	6
درجات اللغة الانكليزية	6	9	6	7	10	8	4	6	9	5

خطوات الحل:

1. نجد مجموع قيم ($\sum x$) الذي يساوي هنا (73).
2. نجد مجموع قيم ($\sum y$) الذي يساوي هنا (70).
3. نستخرج مربع المتغير الأول (x^2) وذلك من خلال تربيع كل قيمة من قيم (x) ثم نجد مجموع هذه التربيعات اي ($\sum x^2$) وهنا يساوي (559).
4. نستخرج مربع المتغير الثاني (y^2) وذلك من خلال تربيع كل قيمة من قيم (y) ثم نجد مجموع هذه التربيعات أي ($\sum y^2$) وهنا يساوي (524).
5. نضرب كل قيمة من قيم ($y \times x$) ثم نجد مجموع حاصل الضرب هذا اي ($\sum y \times x$) هنا يساوي (536).
6. نربع ($\sum x$) ليساوي (5329).
7. نربع ($\sum y$) ليساوي (4900).
8. نقوم بإنشاء جدول, وبالإضافة لعمودي درجات المتغيرين, نضيف ثلاثة أعمدة, (x^2), (y^2), و(x.y). ثم نجمع نواتجها، وكما في الجدول الآتي:

ت	درجات اللغة العربية (x)	درجات اللغة الانكليزية (y)	x^2	y^2	$x.y$
1	7	6	49	36	42
2	8	9	64	81	72
3	5	6	25	36	30
4	7	7	49	49	49
5	9	10	81	100	90
6	10	8	100	64	80
7	5	4	25	16	20
8	7	6	49	36	42
9	9	9	81	81	81
10	6	5	36	25	30
∑	73	70	559	524	536

ثم نطبق معادلة معامل الارتباط بيرسون:

$$r=\frac{n\sum xy-(\sum x)(\sum y)}{\sqrt{[n\sum x^2-(\sum x)^2][n\sum y^2-(\sum y)^2]}}=\frac{(10\times536)-(70\times73)}{\sqrt{[10\times559-(73)^2][10\times524-(70)^2]}}$$

$$=\frac{5360-5110}{\sqrt{[5590-5329][5240-4900]}}=\frac{250}{\sqrt{261\times340}}=\frac{250}{\sqrt{88740}}=\frac{250}{297.89}$$

$r=\underline{0.839\approx0.84}$

إذا، يبلغ معامل الارتباط بين مادتي اللغة العربية واللغة الانكليزية (0.84)

معامل ارتباط الرتب Rank Correlation Coefficient:

هذا المعامل يعرف بمعامل ارتباط سبيرمان (Spearman) أو معامل ارتباط الرتب (رتب القيم وليس القيم الأصلية)، أي، يقيس الارتباط بين متغيرين وصفيين (ترتيبيين)، او أحدهما وصفي (رتبي) والآخر (كمي) ومستمر فاصلي او نسبي (ويتم تحويله الى رتب). وهو أقل دقة من معامل ارتباط بيرسون. ويرمز له بالرمز (r_s)، وقيمته موجبة أقل أو تساوي الواحد الصحيح، وتحسب قيمته وفقا للمعادلة:

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

حيث ان: r_s = معامل ارتباط سبيرمان. 6= عدد ثابت. d^2 = يمثل مجموع مربع الفرق بين رتبتي (x) و(y)، أي، فرق بين رتب القيم لكل زوج من البيانات.

مثال: أراد مرشد تربوي دراسة العلاقة بين ترتيب الولادة للتلاميذ وتحملهم للمسؤولية، وكانت بياناتهم كما في الجدول الاتي:

خطوات الحل:

1. نرتب قيم (تحمل المسؤولية) متغير (X) تنازليا من الاكبر قيمة الى الاصغر، ثم نحدد رتبها، واما ترتيب الولادة المتغير (Y) نرتبها تصاعديا من الاول، ثم نحدد رتبها.
2. نحسب الفرق بين كل رتبتين متجاورتين للمتغيرين، فيكون لدينا العمود (**d**).
3. نربع الفرق، فيكون لدينا العمود (**d^2**).
4. نجمع هذه المربعات (**$\sum d^2$**). وكما في الجدول أدناه:

d^2	**d**	**Rank Y**	**Rank X**	**Y**	**X**
1	**1**	**5**	**6**	**الثالث**	**53**
1	**-1**	**5**	**4**	**الثالث**	**60**
0.25	**-0.5**	**8**	**7.5**	**الرابع**	**50**
1	**1**	**8**	**9**	**الرابع**	**46**
0.25	**0.5**	**1.5**	**2**	**الأول**	**69**
0.25	**-0.5**	**8**	**7.5**	**الرابع**	**50**
25	**5**	**5**	**10**	**الثالث**	**45**
0	**0**	**3**	**3**	**الثاني**	**68**
0.25	**-0.5**	**1.5**	**1**	**الأول**	**72**
25	**-5**	**10**	**5**	**الخامس**	**54**
54	**المجموع** $\sum$				

5. ونطبق القانون:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = r_s = 1 - \frac{6\times 54}{10(100-1)} = 1 - \frac{324}{990} = 1 - 0.33 = 0.67$$

إذا يبلغ معامل الارتباط سبيرمان بين تحمل المسؤولية والترتيب الولادي ($\underline{r_s = 0.67}$)

الفصل السادس

- **الفرضيات**.
- **مستوى الدلالة**.
- **درجة الحرية**.

الفرضيات Hypotheses

الفروض او الفرضيات (Hypotheses)هي توقعات الباحث لنتائج دراسته. وتعد الفروض حلولاً محتملة للمشكلة موضوع الدراسة. وتعتمد صياغة الفروض على النظريات أو البحوث السابقة أو كليهما، كما أنها تستخدم المصطلحات والمتغيرات التي حددها الباحث. والفرض هو حل للمشكلة تؤيده بعض المعلومات أو الحقائق أو الأدلة النظرية أو الدراسات السابقة، ولكن صحته تعتمد على مدى تأييد الأدلة والشواهد والبيانات الفعلية للفرض.

أنواع الفرضيات Kind of Hypotheses:

تنقسم اختبارات الفروض الإحصائية تبعاً لطبيعة المجتمع الإحصائي إلى قسمين:

أولا: اختبارات الفروض الإحصائية المعلمية: وفى هذا القسم يكون معلوم لدينا التوزيع الذي تتبعه البيانات التي لدينا وما إذا كان توزيعا متصلا أم منفصلا (متقطعا) ويكون المطلوب هو اختبار فروض حول معالم المجتمع.

ثانيا: اختبارات الفروض الإحصائية اللا معلمية: في كثير من التجارب والأبحاث يكون لدينا بيانات واقعية يصعب من خلالها التعرف على التوزيع الذي تتبعه ومن هنا نشأت الحاجة إلى ما يعرف باختبارات الفروض اللا معلمية حيث لا تحتاج مثل هذه الاختبارات معرفة شكل التوزيع الذي تتبعه البيانات محل الدراسة، كما يفضل استخدامها عندما يكون حجم العينة المسحوبة من المجتمع صغيرا نسبيا.

ويمكن تقسيم الفرضيات تبعاً لطبيعة الفرضية ذاتها إلى قسمين هما:

أولا: الفرضيات الصفرية Null Hypotheses: وهي الفرضية التي تشير إلى عدم وجود فروق بين المجموعات إذ ان متوسط مجتمع ما على ظاهرة يساوي قيمة محددة، كأن نقول بأن متوسط المقبولين في الجامعة في امتحان الثانوية العامة يساوي (85) أو ان متوسط تحصيل الذكور يساوي متوسط تحصيل الإناث وعادة تصاغ الفرضية الصفرية على النحو الآتي. مثال على ذلك: (لا توجد فروق ذات دلالة إحصائية بين متوسط تحصيل التلاميذ الذين درسوا بطريقة المناقشة والذين درسوا بالطريقة الاعتيادية)، ويمكن التعبير عن هذه الفرضية بالرموز على النحو الآتي:

$$Ho: \mu_1 = \mu_2 \qquad or \qquad \mu_1 - \mu_2 = zero$$

ثانياً: الفرضيات البديلة أو البحثية Alternative or Research Hypotheses: يشير هذا النوع من الفرضيات إلى التنبؤ بالنتائج. إذ يفترض الباحث ان هناك فروقاً بين المجموعات الداخلة في المقارنة. وتنقسم الفرضيات البديلة إلى قسمين:

1- **الفرضية البديلة عديمة الاتجاه** Non-Directional Hypotheses: يشير الباحث في هذا النوع من الفرضيات إلى وجود فروق بين المجموعتين أو أكثر ولكن لا يحدد اتجاه هذا الفرق أي لصالح من. مثال ذلك: يوجد فروق ذات دلالة إحصائية في متوسط المهارات الحركية بين الذكور والإناث ويمكن التعبير عنها بالرموز على النحو الاتي:

$$H_1: \mu_1 \neq \mu_2 \qquad \text{or} \qquad \mu_1 - \mu_2 \neq \text{zero}$$

٢- **الفرضية البديلة المتجهة**: Directional Hypotheses يشير الباحث في هذا النوع من الفرضيات إلى وجود فروق بين المجموعات لصالح مجموعة دون مجموعة أخرى مثال ذلك:

ان متوسط تحصيل الطلبة الذين يدرسون بطريقة المناقشة أعلى من متوسط الذين يدرسون بالطريقة الاعتيادية، ويمكن التعبير عن هذه الفرضية بالرموز على النحو الآتي:

$$H_1: \mu_1 > \mu_2 \qquad \text{or} \qquad \mu_1 - \mu_2 > \text{zero}$$

الأخطاء المتعلقة باختبار الفرضيات:

ان البحث يصمم للإجابة عن سؤال الدراسة، وذلك عن طريق أما دعم أو الفشل في دعم الفرضية البديلة. ولأنه من الصعب التأكد من صحة التحليل الإحصائي، فإن الباحث أحيانا يرفض الفرضية الصفرية على الرغم من أنها صحيحة وهذا يحدث عندما يجد الباحث بيانات في الدراسة تقترح بان هناك فروقا بين المجموعات في الوقت الذي لا توجد فروق حقيقية. وأحيانا أخرى نفشل في إيجاد فروق في الوقت الذي تكون هناك فروقا حقيقية بين المجموعات.

ومن ثم يتضح أن أي حكم أو قرار نتخذه بصدد الفرض الصفري يحتمل الصحة أو الخطأ. ونكون بذلك أمام أربعة احتمالات:

(أ) أن يكون الفرض الصفري صحيحاً، وتأتى النتائج من العينة لا تؤيده، فإننا نرفضه ويكون القرار خاطئاً، والحكم غير صائب ويسمى خطأ ألفا (α) أو نمط "١" ويعنى رفض الفرض الصفري بينما هو في واقع الأمر صحيح.

(ب) أن يكون الفرض الصفري خاطئاً، وتأتى نتائج العينة تثبت صحته، فإننا نقبله ويكون القرار خاطئاً أو الحكم غير صائب ويسمى خطأ بيتا(β) أو نمط "٢" ويعنى قبول الفرض الصفري بينما هو في واقع الأمر خاطئ.

(ج) أن يكون الفرض الصفري صحيحاً، وتأتى نتائج العينة تؤيد صحته فإننا نقبله ويكون القرار سليماً، أو الحكم صائباً، وهذا يدعى بمستوى الثقة ويساوي (1 - α) .

(د) أن يكون الفرض الصفري خاطئاً، وتأتى نتائج العينة تؤيد خطئه فإننا نرفضه ويكون القرار صائباً أو الحكم سليماً (1 -β) <

ويمكن تلخيص الحالات السابقة على النحو التالي:

القرار	طبيعة الفرض الصفري	
	صحيح	خطأ
قبول الفرض الصفري	قرار صائب	خطأ النوع الثاني بيتا(β)
رفض الفرض الصفري	خطأ النوع الأول ألفا(α)	قرار صائب

مستويات الدلالة الإحصائية Level of Significance:

إن القرار الذي يتخذه الباحث فيما يتعلق بالفرض الصفري الذي يود اختباره أو التحقق من صحته يتطلب وجود قاعدة يستند إليها في هذا الشأن. فالباحث يحاول التوصل إلى أدلة من البيانات التي قام بجمعها تمكنه من رفض الفرض الصفري وقبول أو تأييد الفرض البحثي الذي يشتق من إطار نظري يتبناه ويرى انه يفسر الظاهرة تفسيراً منطقياً. لذلك ينبغي أن يحدد الباحث قبل عملية جمع البيانات قيمة احتمالية معينة تبين مقدار الخطأ الذي يقبل أن يقع فيه نتيجة رفضه للفرض الصفري. وبعبارة أخرى إذا قرر الباحث على أساس البيانات التجريبية التي حصل عليها رفض الفرض الصفري، فإن احتمال خطأ هذا القرار يكون أقل من أو مساوياً هذه القيمة التي يطلق عليها مستوى الدلالة الإحصائية أو ألفا. وفيما يأتي جدول بمستويات الدلالة الإحصائية والمتداولة في الدراسات والبحوث التربوية والنفسية:

مستويات الدلالة (الشك في القرار)	0.01	0.02	0.03	0.04	0.05
مستوى الثقة في القرار أو النتائج	0.99	0.98	0.97	0.96	0.95

درجات الحرية Degrees of Freedom**:**

ويقصد بدرجات الحرية (*df*) هي عدد القيم القابلة للتغيير في حساب خاصية إحصائية ما. مثلا، عدد أفراد العينة ناقصاً عدد القيود. فإذا رمزنا لحجم العينة بالرمز (*n*) فإن الحرية في اختيار أفراد العينة هي (*n* - 1) وتسمى بدرجات الحرية. وتختلف وفقاً للاختبارات الإحصائية المستخدمة، أو القيود التي يتم وضعها للمقارنة.

ويعتمد حساب الخصائص الإحصائية المختلفة على مجموعة من المعلومات أو البيانات. ويسمى عدد المعلومات المستقلة عن بعضها والتي تدخل في حساب خاصية إحصائية معينة (كالتباين Variance، والارتباط Correlation ، ...الخ) بدرجات الحرية .(*df*) بشكل عام، عدد درجات الحرية في تقييم خاصية إحصائية معينة يساوي عدد القراءات المستقلة التي تدخل في حساب الخاصية الإحصائية (تباين، ارتباط، ...) ناقص عدد الخصائص الإحصائية المستخدمة في حساب الخاصية الإحصائية المطلوبة (مثل استخدام قيمة المتوسط الحسابي في حساب التباين مثلا. فعلى سبيل المثال: لدينا عينة إحصائية مكونة من (100) قراءة لنتائج امتحان ذكاء أجري على (100) شخص. عدد درجات الحرية في حساب التباين لهذه العينة يساوي عدد القراءات المستقلة "(100)" ناقص عدد الخصائص الإحصائية المستخدمة في حساب التباين وهي هنا تساوي (1) لأننا فقط نستخدم المتوسط الحسابي في حساب التباين. أي أن عدد درجات الحرية لحساب التباين في هذا المثال يساوي (99).

المصدر للمادة الدراسية

وزارة التربية العراقية

كلية التربية المفتوحة

أ.د. أسامة حميد حسن

أ.م.د. صادق عبد النور عزيز

الاحصاء التربوي ٢٠٢٣-٢٠٢٤

التوفيق والسداد الدائمين لكم طلبتي الأعزاء

م.م.خيرالله فرج سبهان الجبوري

الفصل الثاني لسنة 2025