وزارة التعليم العالي والبحث العلمي

جامعة محمد خيضر - بسكرة

معهد علوم وتقنيات النشاطات البدنية والرياضية

محاضرات في مقياس:

الإحصاء الوصفي

موجهة لطلبة السنة ثانية ليسانس

جميع التخصصات

إعداد: د. خـالد دحية

السنة الجامعية:2023-2024

فهرس المحتويات

1. مدخل مفاهيمي لعلم الإحصاء.
2. المتغيرات الاحصائية ومستويات القياس.
3. المعطيات الإحصائية، مصادرها، وطرق جمعها.
4. تنظيم، تصنيف، وتبويب البيانات الإحصائية.
5. العرض الجدولي والبياني للبيانات الإحصائية
6. مقاييس النزعة المركزية – الوسط الحسابي.
7. مقاييس النزعة المركزية – الوسيط.
8. مقاييس النزعة المركزية – المنوال.
9. مقاييس التشتت – المدى
10. مقاييس التشتت –الانحراف المتوسط.
11. مقاييس التشتت– التباين، الانحراف المعياري، ومعامل الاختلاف
12. الارتباط والانحدار الخطي البسيط.

مدخل مفاهيمي لعلم الاحصاء

محتوى المحاضرة:

- **تعريف الإحصاء.**
- **أهمية علم الإحصاء.**
- **علاقة الإحصاء بالعلوم الإنسانية والاجتماعية.**
- **اقسام الإحصاء.**
- **وظائف علم الإحصاء.**

مدخل مفاهيمي لعلم الاحصاء

اشتق لفظ الإحصاء Statistics من الكلمة اللاتينية Stutus والتي تعني منزلة أو مرتبة أو حالة أو وضع"، وعلى وجه التحديد يعتبر الإحصاء من العلوم القديمة والتي ظهرت مع ظهور الانسان على الأرض، كما ظهرت في مواضع كثيرة من القران الكريم لتشير إلى الحصر والعدد.

تستخدم كلمة إحصاء في كثير من الكتابات المتخصصة بمعاني متعددة، فهي قد تعني الإشارة الى مجموعة من القواعد والاجراءات التي تستخدم لاختصار كميات كبيرة من البيانات بحيث تصبح في الإمكان الاستفادة من هذا الاختصار في التوصل إلى استخلاصات محددة يمكن الاسترشاد بها في اتخاذ القرارات.

1. تعريف الإحصاء:

هو "فرع من فروع الرياضيات يشمل النظريات والطرق الموجهة نحو جمع البيانات ووصف البيانات والاستقراء وصنع القرارات". (العدل،2022، ص 15)

يعرف علم الإحصاء بأنه" علم الدولة لكونه يعنى بجمع البيانات الخاصة بالدولة مع تلخيصها وعرضها في كل جداول وروم، فالإحصاء علم يختص بالطرق العلمية لجمع البيانات، وتنظيمها، وعرها، وتحليلها بهدف الوصول القوانين واتخاذ القرارات المناسبة". (رضوان،2002، ص37)

من ناحية أخرى يعرف الاحصاء بانه" علم يدرس مختلف طرق ووسائل جمع البيانات الكمية عن مختلف الظواهر الاقتصادية، الاجتماعية، والطبيعية...، وكيفية تحويلها إلى جداول أو أشكال معينة، لتسهيل فهمها واستخدامها، كما يهدف أيضا إلى تقديم الأساليب المختلفة التي

تستخدم لتحليل هذه المعلومات سواء كانت هذه الأساليب وصفية أو رياضية".(راتول، 2018،ص07)

بشكل عام يقصد بالإحصاء معنيين اثنين:

- الأول كونه عبارة عن مجموعة خاصة من القواعد والاجراءات، وهو المعنى البسيط للإحصاء باعتباره يشير إلى البيانات ذاتها مثل درجات معامل الذكاء ومعدل النقاط في امتحان ما. أي أنه مفهوم العامة وغير المليمين بعلم الإحصاء باعتباره مرادف لكلمة اعداد او بيانات او معلومات أي أن مفهوم الإحصاء هنا يرتبط بالأشخاص والاشياء أكثر منه بالعلوم الأخرى.
- الثاني إمكانية تطبيق تلك القواعد والإجراءات علا البيانات المأخوذة من العينات، أي اعتبار أن الإحصاء علم قائم بذاته يتجاوز جمع وجدولة البيانات واختصارها، إلى كيفية الاستفادة من تلك البيانات، ومن ثم توظيفها في عمليات الاستدلال بحيث يكون لهذه البيانات معاني ودلالات محددة. وهنا يقصد به وجهة النظر العلمية باعتباره علم قائم بجمع وتنظيم وتحليل وتفير البيانات للظاهرة المدروسة لاتخاذ القرار المناسب أي أنه بشكل اكر دقة علم اثبات او نفي حالات عدم اليقين، كما أنه يعمل على الالمام بكل المفردات التي يشملها المجتمع الاحصائي المدروس ومعرفة اوصاف كل مفردة من المفردات –يتم التطرق اليها الحقا بالتفصيل– التي يشملها هذا المجتمع معرفة دقيقة ومحددة بالأعداد ومتوافقة مع الغرض من الدراسة سواء بغرض الجمع العددي او بغرض الدراسات والبحوث المتقدمة والتنبؤات المستقبلية".

2.أهمية علم الإحصاء:

أصبح الإحصاء في وقتنا الحاضر ذا أهمية قصوى، كما أن الاهتمام به يرجع إلى عصور قديمة، إذ اهتمت الحضارات القديمة كالصينية والفرعونية اهتماما خاصا بالإحصاء

لتنظيم الدولة في أوقات السلم والحرب. ليشهد في القرن العشرين أعلى معدلات التطور بخاصة فيما تعلق بتطبيق الطرق والنظريات الإحصائية في مجالات العلوم الاجتماعية والتربوية والنفسية.

تظهر أهمية الإحصاء أكثر في استخدامه في توجيه عمليات جمع البيانات بل تتجاوزها إلى تفسير العلاقات التي تعكسها تلك البيانات، كالعمل عل اجراء مختلف المقارنات بين الكثير من الأشياء وفي مختلف المواقف وبكل مستمر في حياتنا اليومية، هذه المقارنات في جوهرها هي عملية إحصائية بين عديد من الاحتمالات تقترن فعليا بالقياس والتقدير والتقييم.

كما أنه يمكن استغلال الطرق الإحصائية في :

- توفير وصف دقيق لمختلف الظواهر المدروسة.
- يساعد في زيادة دقة الوصف والتي تحتاج دائما إلى اختبار مدى ثبات النتائج المتحصل عليها، فبمجرد الوصول إلى النتائج دون التحقق من ثباتها لا يكفي عادة كأساس يعتمد عليه في تفسير الحقائق وتحقيق الفروض.
- يساعد في تلخيص النتائج المتحصل عليها بشكل يتناسب مع ما هو مطلوب، مع تهيل تسهيل عملية المقارنة وإعطاء صورة اوضح للنتائج وأكثر تبيانا للنتائج المتحصل عليها.
- المساعدة في استخلاص النتائج العامة تبعا للنتائج الجزئية، مع إمكانية تحديد درجة احتمال صحة التعميم المتوصل اليه.
- إمكانية التنبؤ بالنتائج، أي توقع حدوث شيء في المستقبل بناء على معطيات موجودة في الحاضر، أي التنبؤ بالنتائج التي يحتمل الحصول عليها في ظروف خاصة. أي وصع النتائج في شكل كمي يتدل بها مستقبلا.
- اختيار البديل المناسب من بين عدة بدائل متاحة وفق ما توفر من بيانات.

- المساعدة في تلخيص البيانات المتاحة وتقديمها في ابسط وأنسب صورة ممكنة، أي انتقالا من كونها بيانات خام إلى استخلاص حقائق علمية واضحة ومحددة من تلك البيانات، وهذا بعد تصنيفها وتنظيمها وتلخيصها.

3. علاقة الإحصاء بالعلوم الإنسانية والاجتماعية:

يستخدم الإحصاء باعتباره وسيلة مساعدة لجميع الباحثين والمختصين في جميع الميادين العلمية وصانعي القرارات في شتى المجالات، ولا يكاد يخلو ميان من ميدان العلوم المختلفة من أثر استخدام الإحصاء أثرا فعالا وبارزا، كون أنه للحصول عل معلومات ذات قيمة من مختلف البيانات الرقمية فإنه يجب بالضرورة استخدام الأساليب والإجراءات والادوات التي يوفرها لنا الإحصاء.

تأثرت العلوم الإنسانية مثلها مل باقي العلوم بالتطورات التي حققها علم الإحصاء، باعتباره يتبنى نفس خطوات المنهج العلمي في البحث، باحتوائه عمليتين منطقيتين هما القياس والاستنتاج، كما اهم استخدام الأساليب الإحصائية في القيام بالبحوث العلمية بيسر وبوقت وتكلفة أقل.

ظهر اهتمام كبير بتطبيق الطرق الإحصائية المختلفة في مختلف العلوم، بهدف تحليل وتفسير مختلف الظواهر الإنسانية الاجتماعية النفسية، الاقتصادية وغيرها، "وقد صاحب هذا التطور الكبير في النظريات الإحصائية بداية ظهور مجموعة من التخصصات المختلفة تهتم بمجالات واهداف خاصة وقد بلغ هذا التطور قدرا هائلا يكاد يظهرها وكأنها علوم مستقلة، ومن هذه التخصصات بحوث العمليات والاحصاء السكاني، ومراقبة الجودة، والاقتصاد القياسي.

ونظرا لاعتماد العلوم المختلفة على الرياضيات في فهم ظواهرها وقياسها وتفسيرها، فقد افردت لها فروع خاصة تهتم بدراسة ظواهرها باستخدام الاساليب الإحصائية والرياضية ومنها الإحصاء الحيوي والاجتماع الرياضي والقيا الاجتماعي وعلم النفس الرياضي والقياس النفسي والقياس التربوي والاقتصاد الرياضي والتاريخ الاقتصادي الجديد.

كما يمكن القول أن الأساليب الإحصائية المستخدمة في مناهج البحث بصفة عامة تستخدم الان في مجال العلوم الإنسانية بنجاح، وقد أمكن عن طريقها التوصل ال بعض الحقائق العلمية والنظريات.

هناك فروق جوهرية بارزة تبين بوضوح الفرق في تحقيق الأهداف بين الابحاث الطبيعية ونظيرتها في العلوم الإنسانية لما تواجهه من صعوبات:

- ما يميز التفاعلات الاجتماعية المتنوعة عدم خضوعها لنظام آلي مرتب.
- بالنسبة لتتابع الاحداث فإنها تسير باطراد يسهل إيجاد ما يحكم نظمها في صورة قوانين ضابطة لتغيراتها.
- إن السلوك الاجتماعي والعلاقات الإنسانية بين الأفراد في المجتمعات لا تخضع لقالب ولا نظام واحد تُسقط فيه أعمالهم وافكارهم وآراؤهم، وهو ما يمنع التوصل إلى قوانين التنبؤ الاجتماعي.
- تعتمد العلوم الطبيعية وحدات كمية تسهل وصف الظواهر المدروسة وبالتالي تقديمها في صورة عددية تيسر بدورها التنبؤ بها وفق شروط معينة، على خلاف العلوم الإنسانية التي لا تملك وحدات محددة لقياس الظواهر موضوع الدراسة.
- القدرة على التحكم والسيطرة عل العوامل المتعلقة بالظاهرة موضوع الدراسة امر غير مضبوط في العلوم الإنسانية باعتبار صعوبة او عدم استجابة البيئة الاجتماعية موضوع الدراسة للأهداف التي يقصدها الباحث.

4. أقسام الإحصاء:

يمكن تقسيم ميدان علم الإحصاء الة قسمين رئيسيين متداخلين:

أ- **الإحصاء الوصفي:** يستخدم عندما يكون الهدف من البحث وصف الظاهرة التي يدرسها الباحث، كما يسعى الباحث من خلاله ال تلخيص البيانات الرقمية بغية تحويلها من مجرد كم من الأرقام إلى شكل او صورة أخرى يمكن فهمها واستيعابها بمجرد النظر لها. هذه البيانات يمكن تلخيصها والتعبير عنها بأحد اشكال مقاييس النزعة المركزية- سيتم التطرق لها لاحقا بالتفصيل كل واحدة على حدة- وقد يمتد التعبير عنها إلى استخدام أحد مقاييس التشتت، إن ما يميز الإحصاء الوصفي هو وفرة الأساليب الإحصائية التي يممكن استخدامها، مع القدرة على اختزالها جميعا في معلومة او اثنتين، يتم بواسطتها التعبير عن الكل، أي أنه ينطلق من الكل ليصل إلى الجزء من خلال تقديم تلك البيانات على شكل جداول او رسوم بيانية. تقتصر وظيفة الإحصاء الوصفي على وصف العينات فقط، وذلك من خلال البيانات التي يتم جمعها من هذه العينات. يتم وصف هذه البيانات من خلال مجموعة من الأساليب الإحصائية، أهمها:

- الجداول الإحصائية: من أهمها الجداول التكرارية.
- العرض او التمثيل البياني: من أعمدة بيانية، الاشكال الدائرية، المدرج التكراري، المضلع التكراري، المنحنى التكراري، المنحنى التكراري المتجمع.
- النسب والمعدلات.
- مقاييس النزعة المركزية: من متوسط حسابي، الوسيط، المنوال، المتوسط الهندسي، المتوسط التوافقي، الربيعيات، العشيريات، المئينيات.
- مقاييس التشتت: من مدى، الانحراف المتوسط، الانحراف المعياري والتباين، الالتواء، التفلطح
- مقاييس المركز النسبي: تتضمن الرتيبة المئينية، الدرجة المعيارية.

- مقاييس الارتباط:
- مقاييس الانحدار الخطي والتنبؤ.

ب- **الإحصاء الاستدلالي:** يعرف بأنه" مجموعة من الأساليب الفنية التي تستخدم لتفسير قيم النتائج التي يتم التوصل اليها من الأساليب المستخدمة في الاحصاء الوصفي، وذلك بغرض التعرف على خصائص المجتمعات الاصلية استنتاجا من خصائص العينات التي سحبت من هذه المجتمعات" (رضوان، 2002، ص 44). أي أنه من الوظائف المستخدمة في مجال البحث العلمي، ويعمل على اختيار جزء من المجتمع- عينة- وفق الطرق العلمية المعتمدة، لاستخدامها في الوصول إلى نتائج تعمم على باقي مجتمع الدراسة، أي أن الإحصاء الاستدلالي يستخدم البيانات الرسمية التي تم التحصّل عليها من الإحصاء الوصفي للاستدلال بها على حكم او نتيجة ما فيما يختص بالمجتمع الأصلي الذي اشتقت منه العينة.

5- وظائف علم الإحصاء:

يمكن تحديد أهم وظائف علم الإحصاء في الآتي:

أ- **وصف البيانات:** تعتبر طريقة جمع البيانات وتبويبها وتلخيصها من أهم وظائف علم الاحصاء، اذ لا يمكن الاستفادة من البيانات الخام، ووصف الظواهر المختلفة محل الاهتمام، الا إذا تم جمع البيانات وعرضها في شكل جدولي، او بياني من ناحية، وحساب بعض المؤشرات الإحصائية البسيطة التي تدلنا على طبيعة البيانات من ناحية أخرى. وعلى أي حال تحتاج عملية جمع البيانات إلى عملية تسمى القياس والذي يعني " إعطاء تقدير كمي لشيء ما او للشيء المراد قياسه عن طريق مقارنته بوحدة معيارية متفق عليها" (العدل، 2022، ص 48). وباعتبار أن البيانات الخام لا يمكن

الاستفادة منها الا إذا تم تنظيمها بهدف الخروج بملامح عامة، كان يتم تنظيمها في صورة جداول او في صورة رسومات بيانية

ب-**الاستدلال الاحصائي:** يهتم بموضوعين هما:

- التقدير: وفيه يتم حساب مؤشرات من بيانات العينة وتسمى إحصاء، تستخدم كتقدير لمؤشرات المجتمع وتسمى معالم، ويطلق على المقاييس الإحصائية المحسوبة من بيانات العينة في هذه الحالة بالتقدير بنقطة، كما يمكن أيضا استخدام المقاييس الإحصائية المحسوبة من بيانات العينة في تقدير المدى الذي يمكن ان يقع داخله معلمة المجتمع باحتمال معين، ويسمى ذلك التقدير بفترة." (خليل، دس، ص8).
- اختبار الفروض: كل الاختبارات الإحصائية تبدأ بصياغة الفروض العلمية، وتتجه معظم البحوث في التربية البدنية والرياضية إلى استخدام الفرض الصفري وهو فرض يقرر بداية أنه لا يوجد فرق دال بين العينات موضوع الدراسة وهو يتضمن احكاما عما إذا كانت الفروق الظاهرية التي يتم التوصل اليها بين العينات فروقا حقيقية، أم أنها ترجع لأخطاء الصدفة" (رضوان،2002،ص 48)، كما أن القصد من اختبار الفروض هو استغلال مختلف البيانات المتعلقة بالعينة موضوع البحث بهدف الوصول إلى نتائج علمية سليمة يتم من خلالها اتخاذ القرار السليم بخصوص الفروض المحددة حول معالم مجتمع البحث.

ج–**التنبؤ:** فيه يتم استغلال نتائج الاستدلال الاحصائي والتي تعبر عن ظاهرة ما في الماضي، في تحديد ما يمكن أن يحدث لها حاضرا وفي المستقبل.

المتغيرات الاحصائية ومستويات القياس

محتوى المحاضرة

- تصنيف المتغيرات الإحصائية الى متغيرات كيفية وكمية.
- تصنيف المتغيرات الإحصائية الى متغيرات تابعة ومستقلة.
- تصنيف المتغيرات الإحصائية حسب مستويات القياس.
- مقارنة بين مستويات القياس الأربعة.

المتغيرات الاحصائية ومستويات القياس

تنطلق فكرة البحث في أي ميدان من ميادين العلوم المختلفة في ضو ء الاختلافات والتنوع بين الأفراد والظروف، وبالتالي فان الهدف الأساسي لاي بحث عموما هو محاولة فهم وتفسير كيف تتغير الأشياء والكشف عن أسباب التغير. إن مصطلح "متغير يدل على صفة محددة، تأخذ عددا من الحالات او القيم او الخصائص، وتشير البيانات الإحصائية التي يقوم الباحث بجمعها إلى مقدار الشيء او الصفة او الخاصية في العنصر او المفردة او الفرد إلى متغيرات". (العدل، 2002، ص71). أي أن هذ المصطلح يحمل شيئا يتغير ويحمل قيما مختلفة او صفات متعددة، أي أنها توابع ذات قيم متغيرة تتعلق دائما بصفة معينة، وبشكل عام يحمل مصطلح متغير الاشارة إلى سمة او خاصية او صفة تكف عن فروق بغض النظر عما إذا كانت هذه الفروق كمية او نوعية. يمكن تصنيف المتغيرات حسب الغرض من التصنيف، بمعنى أنه يمكن أن تكون كمية أو كيفية، او تصنيفها إلى متغيرات تابعة وأخرى مستقلة، وأيضا من الممكن تصنيفها حب مستويات القياس.

إن وجود شيء ما في الطبيعة يعني أنه يوجد بمقدار، وبالتالي فإنه من الممكن قياسه، فالقياس اجراء يتم في مجالات العلوم المختلفة على أساس وجود الشيء المراد قياسه "وتعتمد عمليات القياس في النشاط الإنساني على أن الظواهر المراد قياسها هي في الواقع أشياء يمكن ادراكها او ملاحظتها او توقع حدوثها، ومن هنا تتأسس فلسفة القياس على أن أي شيء نسلم بوجوده في الطبيعة، فانخ لا بد أن يكون موجودا بمقدار حيث لا يمكن قياس شيء لا وجود له في واقع الحياة العملية، لعدم إمكانية تحديد معالمه ومكوناته تحديدا دقيقا"(رضوان،2002،ص09).

يُشار أنه من الممكن إيجاد صعوبة في قياس شيء ما لعدم توفر وسائل وأدوات قياسه المناسبة، وهو مالا يمنع ملاحظته من خلال مظاهر الأداء الدالة عليه.

1-تصنيف المتغيرات إلى متغيرات كيفية و كمية:

أ- **المتغيرات الكيفية:** هي المتغيرات التي لا يمكن قياسها بالطرق المعروفة، ولا يمكن التعبير عنها ال بقيم عددية صحيحة مثل الجنس، ولون العينين ولون البشرة والعقيدة، يشار إلى أن القيمة هنا لا تعبر عن مقدار الخاصية او السمة عند الفرد بل تعبر عن وجودها او عجم وجودها عنده، أي أن الامر يتعلق بامتلاك الخاصية او عدم امتلاكها، وبالتالي فإنه لا يمكن ترتيب الأفراد طبقا للخاصية- الكيفية- من الأكبر او الأصغر وهكذا.

ب- **المتغيرات الكمية:** هي بيانات يعبر عنها بأرقام عددية تمل القيمة الفعلية لظاهرة ما أي أن المتغير يشير إلى مقدار ما لدى الفرد من خاصية مثل الطول، الوزن، العمر، ...، أي أنها تحمل معنى كميّا وبالتالي يمكن ترتيب الأفراد طبقا لهذه الخاصية- الكمية- من الأكبر إلى الأصغر وهكذا.

2-**تصنف المتغيرات الكمية إلى متغيرات كمية متصلة:** هي عبارة عن سمات او خصائص قابلة للتغير من مشاهدة لأخرى ومن موقف لآخر، وهذه التغيرات تتطلب وضع قيمة لها تقع بين حدّي درجتين معينتين بحيث لا يمكن التعبير عنها بالعدد الصحيح او الكر او كليهما معا، أي أنها تأخذ موجبة او البة صحيحة او كرية مثل درجة الحرارة الاعمار الطول ، المسافة الزمن، السرعة، كما تصنف إلى متغيرات كمية متقطعة تأخذ قيما صحيحة موجبة فقط كمعدل النبض ، عدد لاعبي فريق ما، عدد الأهداف في مباراة ما، كحجم الاسرة مثلا، أي أن القيم فيه تأخذ دائما قيما كاملة تكون في شكل اعداد صحيحة. وعلى هذا فإن خصائص او صفات مثل الجنس ولون العينين،

والجنسية والسلالة...، عبارة عن متغيرات تكشف عن فروق كيفية بين شخص وآخر، بينما خصائص مثل الطول والوزن ...، تكشف عن فروق كمية.

3-تصنيف المتغيرات إلى تابعة ومستقلة:

- **المتغير التابع:** هو المتغير الذي يرغب الباحث من خلال عن تحديد تأثير المتغير المستقل عليه، أي أنه لا يخضع لتحكم الباحث وتتصف بالاتجاه والكمية والمقدار وهي متغيرات عامة مثل السرعة، التحمل، القوة، تشوهات القوام، علامات الاختبارات الخاصة بالأبعاد النفسية والبدنية والمهارية في المجال الرياضي. يشار إلى أن المتغيرات التابعة تتعلق بالسلوك الإنساني بأبعاده الجسمية الصحية الحركية المعرفية الانفعالية الاجتماعية وغيرها، مثل تسجيل أثر الضوضاء على التوتر النفسي الناشئ عن الضغوط، بتسجيل متغيرات من الممكن أن تحدث في ضغط الدم، معدل ضربات القلب، الادرينالين في الدم، والأداء في المهارات الحركية او حتى في امتحان ما باعتبارها متغيرات تابعة.
- **المتغير المستقل:** هي المتغيرات التي من الممكن أن يتم دراتها بالتجريب في البحوث والتجارب العلمية المختلفة، "أي أنه المتغير الذي يتم بحث اثره في متغير آخر ويمكن للباحث التحكم فيه للكف عن تباين هذا الأثر باختلاف قيم او فئات او مستويات ذلك المتغير"(العدل، 2002،ص77). من امثلتها في المجال الرياضي برامج ومناهد التربية البدنية والرياضية في المدارس، أساليب وطرق التدريب، طرق التدريس المختلفة...

4-تصنيف المتغيرات حسب مستويات القياس:

يعرف القياس بأنه:" عملية تقدير للأعداد التي تستمد من ملاحظة اللوك او من الخصائص المميزة للأشياء وفقا لبعض الإجراءات، وفي مجال القيا س في التربية

الرياضية يتم استنتاج الدرجات او الاعداد عن طريق تعيين الأرقام الخاصة بالمشاهدات المختلفة وفقا للعدد من القواعد الدقيقة المعدة مسبقا" (رضوان،2002، ص10). أي أن الأرقام في القياس تستخدم كلغة للوصف للتعبير عن امتلاك رياضي ما لصفة من الصفات وأيضا تعطي قيمة محددة لمدى امتلاكه او حتى قدرة إنجازه لتلك الصفة.

- **مستويات القياس:** على اعتبار أنه من غير الممكن عمليا قياس كل المتغيرات بمقاييس واحدة فإنه هناك اربع طرق او مستويات مختلفة تحدد الدرجات التي تعبر عن خواص الأشياء، كما أنه توجد قواعد مختلفة ومتنوعة يمكن تطبيقها للمستويات الأربع، والتي تعمل–المستويات– على تيسير العمل على اختيار الأسلوب الاحصائي المناسب للبحث.
- **القياس الاسمي:** هو ادنى مستوى من مستويات القياس، ويستخدم في معظم الأحوال مع المتغيرات النوعية ، حيث يتم التصنيف طبقا لخصائص نوعية ، كما و يتم استخدام الأرقام بدل الأسماء، كأن يتم تصنيف الأفراد والاشياء والمجموعات إلى فئات وهذا وفقا لاي متغير او اعتبار كأن يتم التصنيف على أساس السن او الجنس او الدين او لون البرة او نمط الجسم، او البيئة وكلها قياسات اسمية او نوعية. فإذا حددنا مثلا الرقم 1 ليدل على أن المفحوص ذكر والرقم 2 ليدل على أن المفحوص أنثى فإن هذا لا يعني بالضرورة أن 2 أكبر من 1 ، كون أن الرقم هنا ليس له أي معنى كمي ، وإنما يؤدي وظيفة التصنيف فقط وهو ما يبين هنا بشكل واضح أن لا قيمة للأرقام في القياس الاسمي وإنها تحل محل اسما او فئات. يتم التصنيف في القياس الاسمي إلى فئتين او أكثر، على أن يكون هناك تكافؤ بين كل وحدات الفئة ، وأن تتباين فئات التصنيف فيما بينها، بمعنى اكير إذا قمنا بتصنيف ما على أساس الجنس

فإن مجموعة الرجال تضم حتما كل الرجال ومجموعة النساء تضم كل النساء ، أي أن هناك تباين في المتغير والطي هو الجنس من الرجال والنساء، وأيضا وحدات كل فئة متكافئة بالنسبة لمتغير الجنس.

إن ما يميز القياس الاسمي أنه لا يمكن اجراء أيا من العمليات الحسابية المتعلقة بالجمع، الطرح، الضرب، القسمة، كون أن الاعداد التي يتم تحديدها للفئات تكون خاصة بكل فئة من فئات التصنيف كل على حدة،

- من الشروط اللازمة في القياس الاسمي:
- أن تضم قائمة الفئات كل المشاهدات في التجربة.
- أن تتباين الفئات فيما بينها بكل تام.
- الابتعاد عن أي معان تدل على التحيز او التعصب لترتيب فئة على أخرى.

- **القياس الرتبي:** هو المستوى الثاني من مستويات القياس وفيه ترمز الاعداد إلى رتب تبين المواقع النسبية للأفراد او الأشياء، ترتب تصاعديا او تنازليا في صفة او خاصية ما، وهو أبسط استخدام للغة الكم وهو الترتيب باعتبار أن البيانات تدل على وجود ظاهرة معينة وترتيب مجموعة من الأشخاص في هذه الظاهرة، وتعكس مقاييس الرتبة ما إذا كان الشخص او الشيء اصغر او اكبر اثقل او اخف أطول او اقصر اقوى او اضعف بالنسبة للآخرين، وغالبا ما يوضع الأفراد في هذا المستوى في مجموعات وفقا للخصائص البدنية او السمات النفسية لجعل العلاقة متصلة، وتعتبر العبارات الشائعة: اكبر من، أطول من، ا قل من، اذكى من، من امثلة القياس الرتبي او مقاييس الترتيب او الرتبة.

إن الأرقام في هذا المستوى من القياس تُظهر الاختلافات الكمية للخاصية او الظاهرة التي يتم قياسها، فالبيانات الرتبية مجرد رتب تحدد موقع

الشخص في المجموعة دون توفر روط المسافات المتساوية، والرتب نوع من الكم الا أنها لا تدل على اعداد، والفرق بينها وبين الاسمية أن المقاييس الامية تعبر عن عدد دون كم ام القياسات الرتبية فتعبر عن كم دون عدد. فمثلا متسابقو 1500م يرتبون حسب وصولهم إلى خط النهاية الأول، الثاني، الثالث وهكذا. أي أن هذا الترتيب يعكس المهارات والقدرات الفردية للمشاركين في السباق.

وكمثال لترتيب طلبة أحد اقسام السنة أولى جذع مشترك في اختبار مقياس ما وفقا لعلاماتهم الأول، الثاني، الثالث ...الخ، فالطالب أولا حصل على العلامة 18، والطالب الثاني 17، والطالب الثالث 13 من هذا: أن الفرق الحقيقي بين علامات الطالبين الأول والثاني نقطة واحدة، وبين الثاني والثالث هي اربع نقاط ، وبالرغم من أن الفرق بين رتبتي الطالبين الاول والثاني تساوي الفرق بين رتبتي الطالبين الثاني والثالث ، فان الفرق بين العلامات الحقيقية للطالبين الأول والثاني لا تتساوى في الفرق بين علامات الطالبين الثاني والثالث بين العلامات ، أي أنها ليت متساوية البعد عن بعضها البعض حتى وأن كانت متتالية، مع أنه لا يشترط أن تكون متساوية وهو أكثر ما يميز القياس الترتيبي على اعتبار أنه لا يعطي صورة واضحة عن حجم الفروق الموجودة بين الأفراد المتتالية في أي مجتمع.

- **القياس الفتري:** في هذا المستوى تكون فيه الفروق بين المستويات المتتالية متساوية ،أي أن لهذا المتغير وحدة قياس الا أن الصفر او نقطة البداية اختيارية أي أنها نسبية وليست مطلقة ، كما أنها لا تعني غياب الظاهرة او الخاصية المقاسة .

يلاحظ كثرة استخدام هذا المستوى في بحو التربية البدنية والرياضية، وهو يتضمن أرقاما مختلفة تعكس اشياء وظواهر مختلفة اذ يعتمد على ترتيب بعض المشاهدات وفقا لأبعاد معينة.

- **القياس النسبي:** يتميز القياس النسبي بجميع الخصائص السابقة إضافة إلى وجود الصفر المطلق الذي يعني غياب الخاصية من الاطوال والأوزان فمثلا سرعة متسابق في سباق 3000م هي ضعف سرعة متسابق آخر لنفس السباق، كما أنه يشتمل على وحدات قياس متسقة وهي من ادق المقاييس. يوضح الجدول التالي الفروق فيما بين مستويات القياس الأربعة وفقا لأهميتها من الأعلى ال الأسفل.

مقارنة بين مستويات القياس الأربعة

مستوى القياس	الخصائص المميزة	امثلة
القياس النسبي	الأرقام التي يتم الحصول عليها تمثل وحدات متساوية بالنبة للصفر المطلق يمكن اجراء المقارنات المختلفة باستخدام النسب او النسب المئوية	المسافة- الزمن- الوزن- الطول...
القياس الفتري	الأرقام تمثل وحدات قياس متساوية- فئات- وتستخدم المسافات بين المشاهدات لإجراء المقارنات المختلفة	السنة- الحرارة- معامل الذكاء...
القياس الرتبي	الأرقام تدل على الترتيب التدريجي للمشاهدات المختلفة في لتجربة	الترتيب وفقا لاي قواعد او محكات...
القياس الاسمي	الأرقام تمثل فئات تصنيف متباينة الأرقام لا تعبر عن الفروق الكمية في الدرجات الأرقام لا تميز بين المجموعات	النوع- الجنية- اللغة- الدين- المستوى الدراسي...

المعطيات الاحصائية مصادرها، وطرق جمعها

محتوى المحاضرة

- مصادر جمع البيانات الاحصائية.
- أساليب جمع البيانات الإحصائية.
- اجراءات تصميم العينة.
- انواع العينات.
- العينات العشوائية.
- العينات غير العشوائية.
- طرق جمع البيانات الإحصائية.
- أخطاء جمع البيانات الإحصائية.

المعطيات الاحصائية تنظيمها، مصادرها، وطرق جمعها

إن العمل على جمع البيانات الإحصائية من اهم اساسيات العمل الاحصائي، فتوفر البيانات بشكل علمي دقيق وسليم عن الظاهرة المدروسة يعطي نتائج سليمة، وبالتالي إمكانية اتخاذ القرار السليم بناء عليها، يمكن تقسيم البيانات إلى مجموعتين:

- **بيانات وصفية:** هي بيانات غير رقمية، او بيانات رقمية مرتبة في شكل فئات رقمية، تقاس بمعيار اسمي او بمعيار رتبي كما ذكرنا سالف
- **بيانات كمية:** هي بيانات يعبر عنها بأرقام عددية تمثل القيمة الفعلية للظاهرة، تنقسم إلى قسمين هما بيانات فترة وبيانات نسبية كما تم ذكره سالفا.

1.**مصادر جمع البيانات:** بسبب تعدد طبيعة المجتمعات الإحصائية واختلاف البيانات، والامكانيات المتاحة للدراسة، تتعدد تبعا لذلك طرق جمع البيانات. تسمى المنابع التي نأخذ منها البيانات بمصادر جمع البيانات وقد تكون بشكل مباشر او غير مباشر:

أ- **مصادر مباشرة:** يتم جمع البيانات في هذه الحالة بشكل مباشر عن طريق الاتصال والاحتكاك بالوحدات الإحصائية او بالاعتماد على الوثائق التي تكون فيها المعلومات لازالت خاما، وتتميز بالدقة والثقة في البيانات المجموعة، باعتبار الباحث هو بنفسه من يقوم بجمعها من المفردة مباشرة، الا انها تحتاج إلى وقت وجهد كبيرين كما انها مكلفة ماديا.

ب-**مصادر غير مباشرة:** يتم الحصول على البيانات من مصادرها غير الأولية، أي بشكل غير مباشر، حيث تكون مبوبة ومصنفة من طرف اشخاص اخرين او هيئات أخرى عمومية او خاصة او تابعة لمراكز بحث وغيرها، من مزاياها توفير الوقت والجهد، كما أن مستوى الثقة فيها لي بنفس مستوى الثقة مقارنة بالمصادر المباشرة.

2. **أساليب جمع البيانات:** يرتبط أسلوب جمع البيانات المتبع بشكل تام بالهدف من الدراسة وحجم المجتمع المدروس سواء تم استخدام أسلوب الحصر الشامل او أسلوب المعاينة.

أ- **أسلوب الحصر الشامل:** يتم استخدام هذا الاسلوب في حال الرغبة في حصر كل مفردات المجتمع ، أي اخذ كل البيانات مباشرة وبدون استثناء لكل المفردات، من مميزاته توفير الحصول على معلومات دقيقة إذا ما توفرت روط البحث الاحصائي، وهو ما يدعم قوة النتائج المتحصل عليها، الا أنه كثير التكاليف كون عملية الحصر الشامل تتطلب موارد وامكانيات مادية ومالية وبشرية كبيرة، وجهد وتكلفة عالية كما أنه لا يتماشى مع الدراسات التي يلعب فيها الوقت دورا حاسما، وأيضا هناك ارتباط كبير بين حجم المجتمع إذا كان كبيرا واحتمال الخطأ بكل اكبر.

ب-**أسلوب المعاينة:** يتم جمع البيانات في هذا الأسلوب من خلال معاينة جزء من المجتمع محل الدراسة، باختياره وفق طرق علمية سليمة، ودراسته وتعميم نتائج العينة المدروسة على باقي مفردات المجتمع، أي أن العينة المختارة تمثل كل المجتمع الاحصائي تمثيلا دقيقا من حيث المميزات والخصائص، يتميز ب:

- الاقتصاد في الوقت والجهد: على اعتبار أن العمل اقتصر على جزء صغير نسبيا من المجتمع الاحصائي، وهم ما يعني الحاجة إلى وقت

اقل في الاعداد والتحضير، مع امكانية الوصول إلى النتائج في وقت وجهد اقل بكثير مما يحتاجه الحصر الشامل .

- تقليل التكاليف المالية: باعتبار تقليل كل من الوقت والجهد في الاعداد والتحضير فان ذلك يؤدي إلى تقليل النفقات المالية للحصول على البيانات المرغوبة.
- ان استخدام الاساليب الإحصائية العلمية المناسبة يؤدي إلى الوصول إلى بيانات أكثر تفصيلا، وأكثر وضوحا مع التقليل كثيرا من احتمال وقوع الخطأ وعدم قبول النتائج ، وايضا تتيح فرصة للتأكد من مستوى دقة النتائج وبالتالي استخدام نتائجها تلك كتقديرات جيدة لمعالم المجتمع المدروس.
- هناك حالات لا بد فيها من استخدام العينات حصرا، بمعنى أنه لا يمكن استخدام أسلوب الحصر الشامل لجميع المفردات بسبب الاستحالة او الخسائر الكبيرة، مثلا عند القيام بالتحاليل الطبية يلجأ المختبر إلى اخذ عينات فقط من دم الانسان لتحليلها لأن سحب كامل الدم منه تؤدي إلى هلاكه. او اختبار مدى جودة عتاد ما وغيرها.
- لضمان نتائج أكثر دقة عند استخدام أسلوب المعاينة يجب أن تكون العينة المختارة مملة للمجتمع تميلا جيدا لكي لا تكون نتائجه اقل دقة من أسلوب الحصر الشامل.

3. إجراءات تصميم العينة :

يجب أولا تبيين الفرق بين مجتمع الدراسة، والعينة المختارة من المجتمع، كما يجب تبيين مختلف الخصائص المتعلقة بالعينة من تجانس وتماثل وتمثيل.

أ- **المجتمع:** هو مجموعة من المفردات التي تشترك في صفات وخصائص محددة، ومجتمع الدراسة هو الذي يشمل جميع مفردات الدراسة، أي هو الكل الذي نرغب دراسته"(خليل، دس، ص 12).

ب- **العينة:** " هي جزء او شريحة من المجتمع تتضمن خصائص المجتمع الأصلي الذي نرغب في التعرف على خصائصه ويجب أن تكون تلك العينة مملة لجميع مفردات هذا المجتمع تمثيلا صحيحا"(حراوبية،2017،ص15). أي أنها جزء من المجتمع يتم اختياره بطرق مختلفة بغرض دراسة هذ المجتمع.، يمكن أن تكون العينة متجانسة تجانسا تاما بمعنى أن جميع مفردات مجتمع الدراسة متجانسة وتحمل نفس الخصائص، او أن تكون متجانسة تجانسا شبه تامّ ، بمعنى أن هناك تجانس غير تام بين مفردات مجتمع الدراسة، كمت يمكن أن يكون هناك اتفاق في الخصائص بين مجموعتين يريد الباحث دراستهما وهو ما يسمى التماثل بالرغم من صعوبة تحقيقه بنسبة كاملة الا أنه من الواجب جعل المجموعتين قيد الدراسة متماثلتين إلى اكبر قدر ممكن، كما يجب أن تعكس العينة المختارة خصائص مجتمع الدراسة المسحوبة منه أي أن تظهر خصائص مجتمع الدراسة فيها وبنفس ورودها في المجتمع الأصلي. وعموما عند تصميم العينة يجب توفير الإجراءات التالية:

ج-**تحديد حجم العينة:** يراعى فيه اختيار أداة تحديد حجم المجتمع، وطبيعة الخاصية المدروسة وفيما إذا كان تباين المجتمع متوفرا ام لا.

د- **تحديد نوع العينة:** اعتمادا على طبيعة المجتمع الاحصائي المدروس وخصائصه ومقدار التجانس بين مفرداته-كما ذكرنا سالفا-

هـ- **تحديد طريقة اختيار العينة:** والذي يعتمد أساسا على نوع العينة المقرر اختيارها سواء أكانت عشوائية(احتمالية)، أو كانت غير عشوائية (غير احتمالية) ولكل منهما استخداماتها التي تتوقف على الغرض الرئيسي من الدراسة.

4. أنواع العينات:

1.4 –العينات العشوائية: "هي التي يتم اختيار مفرداتها من مجتمع الدراسة بطريقة عشوائية وفق شروط ": (البلداوي، 2007، ص 59)

– كل عينة يمكن اختيارها من المجتمع لها احتمال معلوم، وتبعا لذلك فكل وحدة احتمال معلوم تشمل في العينة، وليس من الضرورة أن يعني هذا الاحتمال المعلوم تساوي الاحتمال لكل وحدة في المجتمع، بل قد يختف وهذا الاختلاف يساعد في حالات المجتمعات غير المتجانسة على توفير دقة أعلى للتقديرات التي نحصل عليها من العينة.

– تسحب العينة باستخدام احدى طرق الاختيار العشوائي، بحيث تتحقق الاحتمالات المعلومة.

– تعتمد الاحتمالات المعلومة عند استخدام نتائج العينات في الحصول على تقديرات جيدة لمعالم المجتمع الذي نقوم بدراسته.

من أنواعها نذكر:

أ– **العينة العشوائية البسيطة:** تستخدم في حال كان المجتمع متجانسا من حيث الغرض او الصفة التي تتعلق بها الدراسة، تؤخذ العينة بشكل يعطي لجميع مفردات المجتمع المراد دراسته نفس الفرصة لأن يكون ضمن العينة، أي أنه لا يجب إعطاء اهتمام لبعض المفردات عن البعض الآخر، وبالتالي ضمان تكافؤ الفرصة في الظهور. أي أن العينة العشوائية البسيطة هي اختيار بسيط يسمح بأن يتحقق لكل مفردة من المجتمع الأصلي درجة احتمال متساوية في الاختيار او الظهور، وأيضا أن يكون اختيار كل مفردة من مفردات العينة بذاتها بصورة مستقلة عن الأخرى. تؤخذ العينة العشوائية بإحدى الطريقتين، اما أن يتم خلط

مفردات العينة بطريقة عشوائية خلطا جيدا، ويتم أخذ مفردات العينة بطريقة عشوائية، واما باستخدام جداول الأرقام العشوائية.

حجم العينة= نسبة العينة* عدد مفردات المجتمع

مثال: إذا اردنا اختيار عينة قوامها 15% من مجموع طلاب قسم سنة أولى جذع مشترك بمعهد علوم وتقنيات النشاطات البدنية والرياضية والبالغ عددهم 300 طالبا.

- يتم حساب عدد أفراد العينة = 100/15 * 300= 45طالبا

- باستخدام جداول الأرقام العشوائية لاستخراج 45 مفردة نتحصل على قائمة متكونة من 45 طالبا، ومن ثم إعطاء رقم متسلسل لكل طالب أي من 1 إلى 45، وبعدها يتم اختيار أي رقم من الأرقام من البداية او المنتصف او في أي مكان وهكذا حتى الوصول إلى الرقم المنشود.، كما يمكن استخدام القرعة وفق أسلوب علمي سليم.

- تجدر الإشارة إلى أن العينة العشوائية البسيطة يمكن أن تسحب بأحد الاسلوبين، إما السحب بدون إعادة، أي رفض الرقم الذي سبق وإن أخذ، او السحب مع الإعادة، أي قبول مرة أخرى الرقم الذي سبق وإن أخذ.

ب- **العينة المنتظمة:** هي نوع من المعاينة العشوائية، يتم فيها الاختيار وفق نظام معين باتباع مسافة ثابتة منتظمة بين كل رقم لمفردة والرقم الذي يليه لمفردة أخرى من المجتمع الأصلي.

مثال: نفس المثال السابق لو أخذنا 15% من عدد الطلاب 300 طالب، يستطيع الباحث أن يختار هؤلاء ء الطلاب بطريقة عشوائية، فيبدأ مثلا بالرقم 3، ثم بعد

كل 10 طلاب يقو م باختيار طالب اخر، أي أنه سيختار الأرقام 3-13-23-33-43... وهكذا، تتميز هذه الطريقة بالقبول بشرط أن يكون اختيار الأرقام من البداية بكل نزاهة ولا يخفي وراءه أي تحيز باختيار طلاب بعينهم.

مثال: إذا ما أردنا التعرف على مدى رضا جمهور كرة اليد عن الجو العام و التنظيم عند الحضور للقاعة الرياضية، باستخدام هذه الطريقة مثلا نقوم بأخذ مكان عند المدخل الرئيسي للقاعة ونستجوب -على سبيل المثال -كل خامس او عاشر فرد يدخل الملعب، او أن نعطي الجمهور أرقاما من 0 إلى N-1 ونختار رقما عشوائيا ضمن هذا المجال، كأن نختار مثلا الرقم 7 بصفة عشوائية ونضيف كل مرة رقما ثابتا بصفة تسلسلية وتكون الارقام المتحصل عليها هي أرقام الجماهير الذين يشكلون العينة.

فإذا اردنا أن يكون حجم العينة مثلا 15 فردا ، نعمل على الاختيار بشكل عشوائي مثلا الرقم 4 والرقم المضاف بانتظام هو10 مثلا، فإن أرقام الجماهير الذين يشكلون العينة هي: 4-24-34-44-54-64-74-84-94-104-114-124-134-144-154 كما تم شرحه في المثال أعلاه.

ج- **العينة الطبقية:** تتيح العينة الطبقية فرصة دراسة كل طبقة من الطبقات على حدة بالإضافة إلى دراسة المجتمع ككل، وتعتمد على تقسيم المجتمع الاحصائي إلى فئات او طبقات، ثم يتم اختيار عينة جزئية من كل طبقة يتناسب حجمها مع حجم الطبقة وتشكل مجموعة العينات الجزئية المختارة ما يسمى بالعينة الطبقية.

مثال: نريد أن نسحب عينة تتكون من 20 طالبا من طلبة احدى كليات جامعة بسكرة والبالغ عددهم 940 طالبا، مقسمون كالتالي:

- طلبة السنة الأولى: 320طالبا
- طلبة السنة الانية:240 طالبا
- طلبة السنة الثالثة: 180 طالبا
- طلبة الأولى ماستر:120 طالبا
- طلبة الثانية ماستر 80 طالبا

نختار من كل طبقة من الطبقات عددا من الأفراد يتناسب مع العدد الكلي للأفراد في تلك الطبقة وفق العلاقة التالية:

عدد افراد العينة من طبقة معينة= عدد افراد العينة× (عدد افراد الطبقة/عدد افراد المجتمع الاحصائي).

تحدد الحالات المطلوبة من كل طبقة عشوائيا او بالطريقة المنتظمة.

د- **العينة العنقودية:** يتم فيها تقسيم مجتمع الدراسة إلى مجموعات جزئية ن ثم تختار من كل مجموعة جزئية مجموعة جزئية أخرى اقل منها تسمى عنقودا، ثم يتم اختيار عينات جزئية من هذه العناقيد بطريقة عشوائية بسيطة، تكل مجموعة العينات الجزئية ما يسمى بالعينة العنقودية.

2.4. **العينات غير العشوائية:** يتم اختيار مفرداتها بطريقة غير عشوائية، أي أنها لا تخضع للعشوائية ويتم اختيارها بانحياز لاعتبارات تحقق للباحث الهدف الفعلي من المعاينة. من اهم أنواعها:

أ. **العينة سهلة المنال:** المعيار الوحيد فيها هو سهولة حصول الباحث على مفردات العينة، تستخدم غالبا في الدراسات الاستطلاعية التي لا تتصف بالدقة الكاملة لكنها تتسم بسرعة

الحصول عليها وقلة التكاليف، وكونها مأخوذة بتحيز فإنها معرضة لكثير من الأخطاء مما يُصعّب تفسير، تحليل، وتعميم نتائجها على باقي مفردات المجتمع.

ب. **العينة الحصصية:** يقوم فيها الباحث بتقسيم المجتمع الاحصائي إلى طبقات، ثم يختار عددا من افراد كل طبقة بحيث يتناسب العدد مع حجم كل طبقة، تشبه العينة العشوائية الطبقية، عند اللجوء الى استخدامها يجب: "(راتول،2018،ص20)

- أن تكون الخصائص الخاصة بالمجتمع والتي تؤثر على موضوع البحث متوفرة.
- أن يكون من الممكن استخدام هذه الخصائص في تقسيم المجتمع الى مجموعات متجانسة كالدخل، السن، الوظيفة ...وغيرها.
- أن تكون هذه الخصائص مؤثرة تأثيرا ملحوظا على موضوع البحث.
- أن يكون عدد هذه الخصائص محدودا حتى لا توجد مجموعات كبيرة يصعب التعامل معها.

ج. **العينة القصدية:** يتم اختيارها باعتبار أن الباحث يعتقد من وجهة نظره أنها تحقق أهداف دراسته.

د. **عينة كرة الثلج:** يلجأ الباحث إلى عينة كرة الثلج في حال كان يجهل الأفراد الذين سيجمع البيانات منهم، فيبدأ باختيار فرد معين، وبناء على استجابته يقرر من سيكون الفرد التالي المختار وهكذا حتى اكتمال العينة.

5-طرق جمع البيانات الاحصائية

تُجمع البيانات الإحصائية حسب طبيعة المجتمع الاحصائي إما بالاتصال المباشر بالعينة الإحصائية، او بالاتصال غير المباشر، او عن طريق الملاحظة المنظمة او التسجيل، او تحليل المحتوى، او الاختبارات والمقاييس.

أ. **طريقة الاتصال المباشر:** في هذه الطريقة يجب الاتصال بشكل مباشر بجميع مفردات المجتمع الاحصائي المرغوب دراسته أو بالعينة المختارة من طرف الباحث بنفسه وهو ما يضمن التحكم في دقة البيانات المتحصل عليها وحتى التحكم في التكاليف ، الوقت، والجهد، كما يمكن الاستعانة بأعوان إحصائيين في حال ما إذا كان مجتمع الدراسة أو العينة المختارة كبيرا، بحيث يستحيل على الباحث القيام بذلك لوحده، تسمح هذه العملية في بتفسير الأسئلة بوضوح كما تسمح بجمع المعلومات في اقصر وقت ممكن، ومن عيوبها ارتفاع التكاليف مع احتمال الوقوع في أخطاء كثيرة إضافة الى احتمال التأثير السلبي للأعوان على مفردات العينة، كما أنها تُعدّ غير صالحة إذا ما كانت الأسئلة تتعلق بأمور شخصية بحتة.

ب. **طريقة الاتصال غير المباشر:** في هذه الطريقة يتم جمع البيانات دون الحاجة للاتصال المباشر بمفردات العينة المختارة، بل يتم الاتصال وفق عدة وسائل كالمراسلة، الهاتف، الانترنت وغيرها... من مزاياها أنها قليلة التكاليف، سرعة الحصول على البيانات، وكذا دقتها، وأيضا إمكانية اجراء مسح لجميع مفردات العينة، ومن عيوبها إمكانية لامبالاة المبحوثين.

ج. **طريقة الملاحظة والرصد:** تحتاج بعض الدراسات إلى متابعة الظاهرة المرغوبة ورصدها عن طريق الملاحظة وقياس جميع تغيّراتها من حركة وزمن، وبالتالي التعرف على محدداتها.

د. **طريقة التسجيل:** هذه الطريقة في الغالب تكون منظمة من طرف هيئات رسمية، وتستخدم خاصة في الدراسات الديمغرافية، من مزاياها ضمان درجة عالية من الدقة في البيانات وتوفير المعلومات بسرعة.

6- أخطاء جمع البيانات:

عند جمع البيانات الإحصائية المتعلقة بالمجتمع الاحصائي او بالعينة المختارة فإنه من الممكن أن ترتكب عدة أخطاء نذكر من بينها:

أ- **أخطا المبحوثين:** يمكن أن تكون متعمدة، ناتجة عن الخوف من المسؤوليات التي قد تترتب على الإجابات الصحيحة، او تكون ناتجة عن عدم الوعي بموضوع وهدف وغاية الدراسة في حد ذاتها، او كأن تخضع لتأثير شخصية ومعتقدات الفهم الذاتي للفرد. ويمكن أن تكون الأخطاء غير متعمدة ناتجة عن عدم فهم السؤال، لأسباب متعلقة بعدم صياغته بشكل سليم، او لسوء تصميم الأسئلة او حتى عدم ملاءمة الوقت.

ب- **أخطاء الباحثين:** تتعلق بشكل تام بالباحث وما يرتكبه من أخطاء تسجيل الإجابات او ترتيبها بما يتناسب والاسئلة المطروحة او حتى بنسيان طرح بعض الأسئلة.

تنظيم، تصنيف، وتبويب البيانات الإحصائية

محتوى المحاضرة

- تنظيم المعطيات الإحصائية.
- تصنيف المعطيات الإحصائية.
- الجداول التكرارية.
- الجدول التكراري البسيط.
- الجدول المزدوج.
- أنواع التكرارات.
- التكرار النسبي.
- التكرار المتجمع.

تنظيم، تصنيف، وتبويب البيانات الاحصائية

يقصد بتنظيم البيانات عرضها في جداول مناسبة حتى نتمكن من تلخيصها وفهمها واستيعابها واستنتاج النتائج منها ومقارنتها بغيرها من البيانات، كما يسهل الرجوع اليها في صورة جداول دون الحاجة الى الاطلاع على الاستمارات الاصلية التي قد يكون عددها كبيرا جدا، ولا يمكن باي حال من الأحوال أن تعطينا نظرة وافية عن نتيجة تلك العملية وأيضا يمكن أن تحمل أسماء أصحابها مما يخل بمبدأ سرية البيانات الإحصائية.

يتم اللجوء الى تنظيم البيانات وتصنيفها وتبويبها بجعلها في صورة مجموعات متجانسة تشترك في صفة او خاصية واحدة او عدة خصائص، أي أن نجعلها في شكل فئات إحصائية مبوبة في صورة مناسبة إما يدويا او آليا، ليتم تفسيرها واستخلاص النتائج منها.
إن تصنيف وتبويب البيانات الإحصائية يتم بوجه عام وفقا لإحدى القواعد التالية:

- تصنيف جغرافي.
- تصنيف تاريخي او زمني.
- تصنيف نوعي او وصفي.
- تصنيف كمي.

1. الجداول التكرارية:

أ- **الجدول التكراري البسيط:** يحتوي على صفة نوعية واحدة فقط، والمقصود به ذلك الجدول الذي يتم فيه وضع قيم الدرجات مرتبة ترتيبا تصاعديا في عموده الأول، اما العمود الثاني فيسمى بعمود التكرار ويسجل فيه عدد مرات تكرار كل درجة او حدث.

مثال: البيانات التالية تمثل علامات حصل عليها 25 طالبا في مقياس علم النفس الرياضي في امتحان السداسي الأول:

17 15 13 15 14 16 17 17 16 11 12 13 14

17 12 15 14 11 10 13 18 16 16 15 14

المطلوب: تبويب هذه البيانات في جدول تكراري بسيط.

الحل:

- نقوم بتحديد أكبر قيمة وأصغر قيمة وترتيبها تصاعديا، فنجد اكبر قيمة 17 اصغر قيمة 10، ثم توضع في العمود الأول من الجدول وتسمى(س) أي أن العمود الأول يشتمل على قيم الدرجات.
- في العمود الثاني نضع عدد مرات التكرار باستخدام العلامات.
- في العمود الثالث نضع التكرار ويرمز له بالرمز (ك)

القيمة(س)	العلامات	التكرارات (ك)
10	/	1
11	//	2
12	///	3
13	///	3
14	////	4
15	////	4
16	////	4
17	////	4
المجموع		20

مثال: أخذت عينة عشوائية من 27 طالبا من طلبة السنة الثالثة ليسانس من معهد علوم وتقنيات النشاطات البدنية والرياضية ليتم استقصاؤهم عن التخصصات التي يدرسونها فكانت الإجابات كما يلي:

تربية حركية	إدارة وتسيير	تربية حركية	تدريب رياضي	إدارة وتسيير	تربية حركية	تدريب رياضي	تدريب رياضي	تربية حركية
تدريب رياضي	تربية حركية	تربية حركية	تدريب رياضي	تربية حركية	تدريب رياضي	تربية حركية	إدارة وتسيير	تدريب رياضي
تربية حركية	تربية حركية	إدارة وتسيير	تربية حركية	تربية حركية	تدريب رياضي	تدريب رياضي	تربية حركية	إدارة وتسيير

المطلوب: وضع هذه البيانات في جدول تكراري بسيط؟

الحل:

التخصص	التكرار
تربية حركية	13
تدريب رياضي	05
إدارة وتسيير رياضي	09
المجموع	27

ب- **الجدول المزدوج:** يتم التصنيف فيها على اساس صفتين نوعيتين في الوقت نفسه في الوحدة الإحصائية.

مثال: الجدول التالي يوضح مدى ممارسة طلبة معهد علوم وتقنيات النشاطات البدنية والرياضية للرياضة خارج أوقات الدراسة حب الجنس لعينة من 28 طالبا وطالبة.

الرقم	الجنس	الرد		الرقم	الجنس	الرد		الرقم	الجنس	الرد		الرقم	الجنس	الرد
01	ذ	نعم		08	ذ	نعم		15	ذ	نعم		22	ذ	نعم
02	أ	نعم		09	أ	لا		16	ذ	نعم		23	أ	لا
03	ذ	لا		10	ذ	نعم		17	ذ	نعم		24	ذ	نعم
04	ذ	نعم		11	أ	نعم		18	ذ	نعم		25	ذ	نعم
05	أ	نعم		12	ذ	نعم		19	أ	نعم		26	ذ	نعم
06	ذ	نعم		13	ذ	نعم		20	ذ	نعم		27	ذ	نعم
07	أ	لا		14	أ	نعم		21	أ	لا		28	ذ	نعم

الجنس الممارسة	يمارس	لا يمارس	المجموع
ذكور	18	01	19
إناث	05	04	09
المجموع	23	05	28

2- أنواع التكرارات:

أ- **التكرار النسبي:** يحسب التكرار النسبي بقسمة تكرار المجموعة على مجموع التكرارات من خلال المعادلة:

$$\text{التكرار النسبي} = \frac{\text{تكرار المجموعة}}{\text{مجموع التكرارات}}$$

حيث أن:

- التكرار النسبي يرمز له بالرمز: fi
- تكرار المجموعة يرمز له بالرمز: n_i
- مجموع التكرارات يرمز له بالرمز: N
- إن مجموع ا لتكرارات النسبية يساوي الوحدة $\sum fi = 1$

ب- **التكرار النسبي المئوي:** لمعرفة نسبة تكرار أي فئة من مجموع التكرارات يتم إيجاد التكرار النسبي وذلك بقسمة تكرار كل فئة على مجموع التكرارات، كما ذكرنا في الأعلى، فالتكرار النسبي للفئة اذن هو نسبة تكرار تلك الفئة الى مجموع التكرارات وغالبا ما يتم ضربه في 100 للحصول على التكرار النسبي المئوي، ويرمز له بالرمز f_i % وهو يعطي النسبة المئوية لتكرار كل فئة من مجموع التكرارات ومجموع التكرارات النسبية المئوية يساوي 100.أي: $f_i \ \% = \frac{ni}{N} \times 100$

مثال: الجدول التالي يظهر توزيع عينة من الطلبة حسب النتائج المتحصل عليها في فوج دراسي لتخصص السباحة يتكون من 10 طلبة.

المطلوب: إيجاد التكرار المئوي؟

I	العلامة	عدد الطلبة
1	5	1
2	6	4
3	7	2
4	8	1
5	9	1
6	10	1
المجموع		10

بتطبيق المعادلة :$f_1 \ \% = \frac{1}{10} \times 100 = 10$ %

هذا يعني أن 10% من الطلبة حصلوا على علامة تقدر ب5 ، ويتم إيجاد بقية التكرارات المئوية بنفس الطريقة، وتظهر باقي النتائج كالتالي:

f_i %	f_i	العلامة	I
10	1	5	1
40	4	6	2
20	2	7	3
10	1	8	4
10	1	9	5
10	1	10	6
100	10		المجموع

ج- **التكرار المتجمع** : هو مجموع التكرارات المقابلة لجميع القيم الأقل من الحد الأعلى الحقيقي لفئة معينة، ويتم حسابه إذا اردنا حساب عدد المشاهدات او التكرارات التي تساوي او تقل عن قيمة معينة.

- **التكرار المتجمع الصاعد:** يقصد به تجميع تكرار كل فئة على جميع التكرارات السابقة لها بحيث يكون مجموع التكرار التصاعدي للفئة الأخيرة مساو لمجموع التكرارات. ويرمز للتكرار البسيط الأول ب n_1، والتكرار التجميعي الصاعد للقيمة او الفئة الثانية هو $n_1 + n_2$ ، والتكرار التجميعي الصاعد للفئة الأخيرة يساوي $\sum ni$
- **التكرار المتجمع النازل:** يقصد به تجميع تكرار كل فئة على جميع التكرارات التالية لها بحيث يكون مجموع التكرار التنازلي للفئة الأولى مساو لمجموع التكرارات. أي أن التكرار التجميعي النازل للفئة الأولى يساوي $\sum ni$ والتكرار التجميعي النازل للفئة الثانية هو $\sum ni - n1$ ، اما التكرار التجميعي النازل للفئة الأخيرة يساوي التكرار المطلق للقيمة او الفئة الأخيرة.

مثال: تم تطبيق اختبار حركي لطلبة فوج ما مكون من 15 طالبا وكانت النتائج كالتالي: **8 11 5 9 12 8 7 7 9 10 11 12 12 6 8**

الدرجة	التكرارات	التكرار المتجمع الصاعد	التكرار المتجمع النازل	التكرار المتجمع الصاعد النسبي	التكرار المتجمع الصاعد المئوي
5	1	1	15	0.067	6.70
6	1	2	12	0.133	13.30
7	2	4	10	0.266	26.60
8	3	7	9	0.466	46.60
9	2	9	7	0.600	60.00
10	1	10	4	0.660	66.00
11	2	12	3	.0.800	80.00
12	3	15	2	1.000	100.00
مجموع ك=15					

العرض الجدولي والبياني للبيانات الاحصائية

يمكن عرض البيانات في صورة جدول تكراري، ويختلف كل الجدول طبقا لنوع البيانات وحسب عدد المتغيرات.

1. عرض البيانات جدوليا:

مثال: تمثل البيانات التالية تقديرات 36 طالبا من طلبة قسم التدريب في مقياس العاب القوى.

جيد	ممتاز	جيد	مقبول	جيد	جيد	ممتاز	ممتاز	جيد
حسن	جيد	حسن	مقبول	مقبول	مقبول	مقبول	جيد	مقبول
جيد	حسن	جيد	حسن	جيد	ممتاز	حسن	جيد	حسن
جيد	جيد	حسن	جيد	جيد	جيد	مقبول	مقبول	جيد

المطلوب: اعرض البيانات في شكل جدول تكراري؟

الحل: الجدول التكراري

يتكون الجدول التكراري من عمودين واحد يمثل صفات المتغير والأخر يمثل التكرارات

التقدير	التكرار
ممتاز	04
حسن	07
جيد	17
مقبول	08

مثال: تمثل البيانات في الجدول أدناه علامات مقياس الإحصاء ل 60 طالبا من قسم لطلبة قسم التدريب.

15	**12**	**17**	**13**	**15**	**13**	**10**	**12**	**10**	**13**
10	**13**	**10**	**11**	**11**	**11**	**14**	**16**	**14**	**15**
14	**18**	**11**	**12**	**13**	**15**	**17**	**15**	**17**	**15**
13	**14**	**13**	**12**	**14**	**12**	**17**	**16**	**14**	**13**
10	**12**	**12**	**17**	**14**	**17**	**15**	**15**	**11**	**17**
16	**11**	**14**	**14**	**18**	**13**	**10**	**11**	**12**	**11**

المطلوب:

- اعداد جدول توزيع تكراري؟

الحل:

- حساب المدى: المدى= أكبر قيمة–أصغر قيمة= 18−08=8
- عدد الفئات: نفترض أن عدد الفئات هو 4
- حساب طول الفئة:

C= $\frac{E}{R}$ = $\frac{8}{4}$=2

- الفئة تبدأ بقيمة تسمى الحد الأدنى، وتنتهي بقيمة تسمى الحد الأعلى:
- الحد الأدنى للفئة الأولى هو اقل درجة قيمة أي ان الحد الأدنى للفئة الأولى = 10
- الحد الأعلى للفئة الأولى = الحد الأدنى+ طول الفئة

2+10=

13=

- إذن الفئة الأولى هي 10 الى أقل من 12
- الحد الأدنى للفئة الثانية = الحد الأعلى للفئة الأولى= 12
- بنفس طريقة تحديد الفئة الأولى يتم تكوين حدود الفئات الأخرى، وبالتالي يكون الجدول التكراري:

فئات الدرجة	التكرارات	التكرار النسبي المئوي
10-	6	10.00
12-	17	28.33
14-	18	30.00
16-18	19	31.66
المجموع	60	100.00

2. العرض البياني للبيانات الإحصائية:

تساهم الرسوم البيانية في إعطاء انطباع أفضل وتبين بشكل أكثر وضوحا الخصائص الرئيسية للبينات المقدمة، كما تساعد في تسهيل قراءتها بعيدا عن الجداول التي يمكن ان تكون مكتظة بالمعلومات، يمكن ان تظهر هذه البيانات الإحصائية في عدة اشكال بيانية كالأعمدة، قطاعات دائرية، مدرجات تكرارية، مضلعات تكرارية، منحنيات تكرارية.

- العرض البياني للبيانات النوعية: يمكن عرض هذه البيانات في كل دائرة بيانية او أعمدة بيانية، من خلال وصف ومقارنة مجموعات أو مستويات هذا المتغير.

الدائرة البيانية: يتم رم دائرة وتقسيمها الى قطاعات دائرية تتناسب مساحة كل قطاع مع تكرار الفئة التي تمثلها، فالفئة الأكثر تكرارا تقابل القطاع الأكبر مساحة والفئة الأقل تكرارا تقابل

القطاع الأقل مساحة، أي ان يتم ذلك بتقسيم الكل الى أجزاء وكل جزء يمثل بقطاع دائري بحيث ان زاوية راس كل قطاع دائري تعطى حسب المعادلة التالية:

قياس الزاوية= 360^0 * التكرار النسبي للمجموعة

مثال: الجدول التكراري الاتي يبين توزيع عينة حجمها 1000 طالب حسب الجامعة التي ينتمي اليها.

الجامعة	بسكرة	البويرة	مستغانم	باتنة	المسيلة	الجزائر	ورقلة
العدد	140	120	110	140	160	230	100

المطلوب: مثل البيانات أعلاه في شكل دائرة بيانية؟

الحل:

- تحديد مقدار الاوية المخصصة لكل منطقة، بتطبيق المعادلة:

التكرار النسبي المطلق* 360^0 = مقدار الاوية المخصص للجامعة

الجامعة	عدد الطلبة	التكرار النسبي	مقدار الزاوية
بسكرة	160	0.16	57.6^0=0.16*360
البويرة	140	0.14	=0.14*360 50.4^0

مستغانم	110	0.11	$39.6^0=0.11*360$
باتنة	120	0.12	$43.2^0=0.12*360$
المسيلة	140	0.14	$=0.14*360$ 50.4^0
الجزائر	230	0.23	$82.8^0=0.23*360$
ورقلة	100	0.10	$36^0=0.10*360$
المجموع	1000	1.00	360^0

يتم رسم الدائرة وتقسيمها الى اجزاء تتنسب مع مقدار الزاوية المخصصة لكل جامعة. من الكل أعلاه يلاحظ ان نسبة الطلبة التي تنتمي مثلا الى جامعة بسكرة تبلغ 16 %بينما بلغت نسبتهم في الجزائر العاصمة 23.%

يمكن اعتبار الاشكال الدائرية من أكثر الاشكال استخداما عند عرض البيانات الجدولية، حيث يتم تقسيم الدائرة الى اجاء يدل كل جزء على نسبة معينة من البيانات الكلية كما ذكرنا سلفا.

تستخدم الدائرة في تمثيل البيانات في حالة البيانات التي يمكن تصنيفها الى فئات وفقا للصفات المميزة لها، ومن ثم فهي تصلح في حالة القياس الاسمي حيث تمثل الاعداد الأرقام فئات تصنيف متباينة كالنوع والجنسية والديانة واللغة والمستوى الدراسي وغيرها، ويشترط ان يتراوح عدد من هذه الفئات من 3 الى 5 فئات.

تمثل مساحة الدائرة القيمة الكلية للبيانات أي نسبة 100% ثم تقسم الى قطاعات كل قطاع منها يمثل قيمة او نسبة كل متغير من المتغيرات وفقا لعدد تكراراته في الظاهرة ككل.

مثال: يتمن برنامج الاعمال التطبيقية لطلبة التدريب الرياضي بمعهد الرياضة خمس فئات كبيرة من الأنشطة هي الأنشطة الجماعية، والألعاب الفردية والزوجية، والعاب القوى، والجمباز، وأنشطة اللياقة البدنية، وقد خصص لتدريس البرنامج 36 أسبوعا خلال الموسم الجامعي موزعة على الأنشطة التالية:

- أنشطة الألعاب الجماعية 16اسبوعا
- أنشطة الالعاب الفردية والزوجية 10 أسابيع
- أنشطة العاب القوى 6 أسابيع
- أنشطة الجمباز اسبوعان
- انشطة اللياقة البدنية اسبوعان
- المجموع: 36 أسبوعا

المطلوب: تمثيل هذه البيانات باستخدام الشكل الدائري.

- لتمثيل هذه المجموعة من البيانات باستخدام الكل الدائري، نقوم برسم دائرة الزاوية المركزية لها تساوي 360^0.
- نقوم بتقسيم الزاوية 360^0 على عدد الأسابيع كالتالي:
- $360^0 /36=10$
- نقوم بحساب الزوايا المركزية المخصصة لكل نشاط كالتالي:
- الأنشطة الجماعية = $16*10= 160^0$
- الانشطة الفردية والزوجية= $10*10=100^0$

- العاب القوى: 6*10= 60 ^{0}
- الجمباز: 2*10= 20 ^{0}
- اللياقة البدنية: 2*10= 20 ^{0}
- المجموع 360 ^{0}
- نقوم برسم دائرة مقسمة الى خمس قطاعات مع تبيين الزوايا المركزية والنسب المئوية المخصصة لكل نشاط من الانشطة الخمسة.
- ان النسب المئوية للزمن المخصص لكل نشاط من الأنشطة الخمسة هو:

الأنشطة الجماعية: 16*36/100 = 44.44

الانشطة الفردية والزوجية: 10*36/100 = 28.87

العاب القوى: 6*36/100 = 16.67

الجمباز: 2*36/100 = 5.56

اللياقة البدنية 2*10036= 5.56

الاعمدة البيانية التكرارية: هي رسوم بيانية ملائمة لتوزيع متغير كمي منقطع او نوعي، وهي عبارة عن عدد من الاعمدة بحيث يتم تمثيل الفئات افقيا وتمثل التكرارات راسيا، يرسم عمود واحد لكل فئة بحيث يكون ارتفاع كل عمود يمثل التكرارات او التكرار النسبي المئوي المرتبط بكل فئة في الجدول التكراري.

مثال: البيانات التالية تمثل التوزيع التكراري البسيط لعلامات 40 طالبا في امتحان مقياس العاب القوى لفوجين اثنين من معهد الرياضة.

الفوج الأول				الفوج الثاني			
التقدير	التكرار	التكرار النسبي	التكرار المئوي	التقدير	التكرار	التكرار النسبي	التكرار المئوي
ممتاز	5	0.125	12.5	ممتاز	3	0.1	10
جيد جدا	8	0.2	20	جيد جدا	6	0.2	20
جيد	16	0.4	40	جيد	10	0.33	33
مقبول	5	0.125	12.5	مقبول	6	0.2	20
راسب	6	0.15	15	راسب	5	0.17	17
المجموع	40	1	100	المجموع	30	1	100

مثال:

البيانات التالية تمثل التوزيع التكراري لبيانات مجمعة لأطوال 50 طالبا مقدرة بالسنتيمتر.

فئات الطول(سم)	-145 145.9	-155 165.9	-165 174.9	-175 184.9	185-185 9.
التكرارات	3	9	21	13	4

Commenté [dk1]:

المطلوب: تمثيل هذه المجموعات من البيانات باستخدام الاعمدة البيانية؟

الحل: يتم تمثيل هذه البيانات باستخدام الأعمدة البيانية من خلال:

- تخصيص المحور الافقي لتمثيل الفئات مادامت الفئات متساوية الطول 9.9 سم أي حوالي 10 سم.
- يتم تخصيص المحور العمودي للتكرارات حيث ان اكبر تكرار في الجدول هو 21.
- نقوم بتويع الدرجات التي تمثل الفئات على المحور السيني بحيث تمثل كل فئة بمستطيل حدود قاعدته هي الحدود الحقيقية لتلك الفئة وارتفاعه يتناسب مع تكرار الفئة مع ملاحظة ان نبدأ بأصغر فئة .
- يلاحظ ان التكرارات تمثل بمستطيل يغطي الفئة كلها ويكون ارتفاعه معبرا عن ارتفاع الفئة.
- يفترض من المدرج التكراري ان التكرار موزع بانتظام على جميع قيم الفئة بعكس المضلع التكراري والذي يفترض فيه ان جميع قيم الفئة تمثلها قيمة واحدة هي مركز الفئة.

المضلع التكراري:

هو عبارة عن شكل مغلق نحصل عليه من توصيل التكرارات المقابلة لمراكز الفئات على المحور الافقي بخطوط مستقيمة تعطينا في النهاية خطا منكسرا مغلقا من طرفيه الأيمن والايسر.

- يتكون الرسم البياني الاحصائي من محورين كما ذكرنا سالفا، الافقي للفئات والعمودي للتكرارات.
- نحسب عدد الفئات ، ومن ثم تحديد مقياس رسم مناسب لتميل وحدات البيانات المعطاة على كل من المحورين.
- وضع نقطة فوق النقطة المنصفة للفئة تكون مقابلة لتكرار الفئة.
- توصيل النقاط المتتالية بخطوط مستقيمة تعطينا مضلع تكراري ص 99 .

مثال: يمثل الجدول التالي بيانات لعلامات 150 طالبا في امتحان القدرة الحركية، كما يبين التوزيعات التكرارية والحدود المعلنة والحقيقية ومنتصف الفئات للتوزيعات المتعلقة بعلامات الطلبة.

الفئات	الحدود الحقيقية للفئات	التكرارات	مراكز الفئات	التكرار المتجمع الصاعد
34-30	34.5-29.5	4	32	4
39-35	39.5-34.5	6	37	10
44-40	44.5-39.5	7	42	17
49-45	49.5-44.5	8	47	25
54-50	54.5-49.5	11	52	36
59-55	59.5-54.5	12	57	48
64-60	64.5-59.5	10	62	58
69-65	69.5-64.5	17	67	75
74-70	74.5-69.5	23	72	97
79-75	79.5-74.5	20	77	118
84-80	84.5-79.5	13	82	131
89-85	89.5-84.5	9	87	140
94-90	94.5-89.5	7	92	147
99-95	95.5-94.5	3	97	150

- ان مقياس الرسم على المحور الافقي هو 1 سم لكل 5 درجات من درجات القدرة الحركية.
- كما ان مقياس الرسم على المحور العمودي هو 1 سم لكل تكرارين.
- تم تمثيل مراك الفئات بنقاط منتظمة متساوية المسافة تأسيسا على ان طول الفئة يساوي5
- تم إضافة فئتين أحدهما اعلى التوزيع والأخرى افل التوزيع تكرار كل منهما يساوي صفرا كالتالي:

الفئات	التكرار	مركز الفئة
29-26	صفر	27 =2/29+26
34-30	04	
99-95	3	
104-100	صفر	27=2/104+100

وقد قمنا بتمثيل هاتين الفئتين بنقاط تمثل مراكز الفئات على المحور الافقي وهي النقطتان102،27، حيث تكرار كل فئة من هاتين الفئتين يساوي صفرا.

- في المضلع التكراري يفترض ان جميع قيم الفئة تمثلهم قيمة واحدة هي مركز الفئة، حيث يمكن اعتبار ان جميع قيم المفردات الواقعة داخل كل فئة تساوي تقريبا مركز الفئة، وهذا يعني اننا نعتبر القيمة 32 تكررت 4 مرات، والقيمة 37 تكررت 6 مرات، والقيمة 42 تكررت 7 مرات، وهذا بالنسبة لبقية الفئات.
- يمكن اعتبار الفئتين(26-29، 100-104) فئتين متطرفتين واحدة اقصى اليمين والثانية اقصى اليسار وتكرار كل منهما يساوي صفرا.

- من مميزات المضلع التكراري انه يمكن رسم تويعين تكراريين لعلامات مجموعتين من الافراد في متغير واحد، وذلك بغر عقد مقارنة بين المجموعتين عن طريق ملاحظة اختلاف التويعين، حيث يتطلب ذلك استخراج النسب المئوية للتكرار في كل فئة بالنبة لمجموع التكرارات في كل مجموعة على حدة، ويستخدم هذا الاجراء لتوحيد مجموع التكرارين في المجموعتين يجعله يساوي100.
- **المضلع التكراري المتجمع:**

يتم الحصول على المضلع التكراري المتجمع من مجموعة البيانات المبينة في الجول الخاص بالمثال السابق وبالشكل التالي:

- تمثيل الحدد الأعلى الحقيقي للفئات بنقاط على المحور الافقي.
- تمثيل التكرار المتجمع الصاعد بنقاط على المحور العمودي.
- رصد التكرار المتجمع للفئة فوق الحد الأعلى الحقيقي لها، ثم نقوم بتوصيل النقاط الناتجة بالنسبة لكل فئتين بخط مستقيم فنحصل على المضلع التكراري المتجمع
- تم استبدال التكرارات بالتكرار المتجمع الصاعد على المحور العمودي.
- تم تمثيل الحد الأعلى للفئات (14 فئة) على المحور الافقي مبتدئين بالحد الأعلى للفئة الأولى 34.5 حتى الحد الأعلى للفئة الاخيرة95.5.

الفرق بين المضلع التكراري والمضلع التكراري المتجمع:

- في المضلع التكراري تستخدم مراك الفئات كقيم على المحور الافقي.
- في حين تستخدم الحدود العليا للفئات في المضلع التكراري المتجمع كقيم على المحور الافقي.
- في المضلع التكراري تستخدم تكرارات الفئات على المحور العمودي.

- في حين يستخدم التكرار المتجمع في المضلع التكراري المتجمع كقيم على المحور العمودي.

المنحنى التكراري:

يتم الحصول عليه عموما بتقسيم المحورين الافقي والعمودي وتعيين مواقع النقط كما في حالة المضلع التكراري تماما، ثم نرسم منحنى ممهد باليد يمر بأكبر عدد ممكن من هذه النقط ويمر بتوازن خلال هذه النقط، ويتم اقفاله كما في حالة المضلع التكراري السابق، ويمكن التخلص من فكرة التمهيد باليد لعدم دقتها باستخدام فكرة أخرى تسمى المتوسطات المتحركة.

أي انه يتم رسمه بنفس الطريقة المتعلقة برسم المضلع التكراري مع استعمال الخطوط المنحنية بدلا من الخطوط المستقيمة المنكسرة ، أي انه ينتج من المضلع التكراري عن طريق جعله منحنى بدلا من خطوط منكسرة.

يتم استخدامه لإعطاء صورة عن كل التوزيع بوجه عام، دون النظر الى التفاصيل التي قد تبدو في عدم انتظام شكل التوزيع، كما ان رسمه يتم بصورة اجتهادية كما يمكننا اعتبار انه يمر بأكبر عدد من النقاط الممثلة بتكرارات الحقيقية ويشترط ان يقترب المنحنى بقدر الإمكان من النقاط التي يمر بها، ولهذا فان المنحنى التكراري يسمى أحيانا بالمدرج التكراري.

مثال:

الجدول التكراري التالي يمثل توزيع نتائج الاختبارات لمفهوم الذات لعدد من الطلاب قوامهم 52 طالبا .

المطلوب:

ارسم المنحنى التكراري للتوزيع الأصلي بالتمهيد باليد ثم استخدام المتوسطات المتحركة.

التوزيع الاصلي		التوزيع باستخدام المتوسطات المتحركة		
الفئات	ك	الفئات	ك	المتوسط المتحرك
		104-100	صفر	
109-105	3	109-105	3	3.67
114-110	8	114-110	8	5
119-115	4	119-115	4	6.33
124-120	7	124-120	7	7
129-125	10	129-125	10	9.67
134-130	12	134-130	12	9.33
139-135	6	139-135	6	6.67
144-140	2	144 -140	2	2.67
		149-145	صفر	
المجموع	52	المجموع	52	

الطريقة المتبعة لاستخراج المتوسطات المتحركة:

يعطى لكل فئة تكرار يعادل متوسط تكرارها مضافا اليه تكرار الفئة السابقة لها وتكرار الفئة اللاحقة لها، أي ان:

$$\text{المتوسط المتحرك للفئة الأولى} = \frac{\text{تكرار الفئة الاولى تكرار الفئة الابقة لها+تكرار الفئة اللاحقة لها}}{3}$$

وبما انه لا توجد فئة قبلها فيكون تكرار الفئة السابقة لها = صفر

وعليه:

- فان المتوسط المتحرك للفئة الأولى= 3.67
- المتوسط المتحرك للفئة الثانية = 5

اشكال المنحنيات التكرارية:

تأخذ المنحنيات التكرارية اشكالا متعددة وتختلف باختلاف البيانات المعطاة ومن أنواعها :

- **المنحنى الطبيعي او الاعتدالي:** يسمى أيضا بالمنحنى الجرسي نظرا لكونه يشبه الجرس او الناقوسي او الطبيعي ويلاحظ تطابق نصفيه تماما وعندما تأخذ قمة المنحنى شكلا مسطحا وتختفي القمة يصبح المنحنى مفرطحا وحينما ترتفع القمة ويصبح الجزء المحيط بالقمة نحيلا يسمى المنحنى مدببا او نحيل القمة
- **المنحنى الملتوي:** يمكن ان يكون منحنى ملتوي التواء موجبا وآخر ملتوي التواء سالبا ويلاحظ ان القمة لا يمكن ان يرسم عندها محورا للتماثل بحيث يصبح نصفي المنحنى غير متماثلين.
- **المنحنى المنوالي:** وقد يكون للمنحنى أكثر من قمة ويسمى المنحنى في هذه الحالة منحنى متعدد القيم او منوالي.
- **التوزيع المستطيل:** يظهر المنحنى بهذا الكل حينما تتساوى التكرارات الخاصة بالفئات المختلفة.
- **الاغصان والأوراق:** التوصل الى شكل بياني بطريقة الأغصان والأوراق يمكن اعتباره مرادفا للمدرج التكراري، ويمكن التعامل به اذا كان عدد المشاهدات او القيم في حدود 100 قيمة، فالأغصان والأوراق شانها شان المدرج التكراري

تفيد في تبيين موع تمرك المشاهدات وتفيد في ابراز مدى التماثل(الاقتراب، الاعتدالية، ...) او عدم التماثل (الالتواء الى اليمين او الالتواء الى اليسار..).

- يتميز العرض البياني بفكرة الاغصان والأوراق عن المدرج التكراري في الاحتفاظ بكل المعلومات المتعلقة بقيمة كل مشاهدة، بحيث يمكن قراءة القيمة الفعلية لكل مشاهدة وهو ما لا يسمح به شكل المدرج التكراري.

الغصن عبارة عن عمود يقع في الجانب الاير من الكل، اما الأوراق فتقع الى اليمين فهي لسلة من الأرقام قد يختلف عددها من سطر الى اخر.

مثال:

نفترض انه لدينا بيانات خاصة بالإصابات الرياضية لدى عينة من طلاب قسم التدريب في معهد الرياضة.

-102-113-83-90-100-112-78-82-95-100-110-125
-95-93-93-106-106-119-117-114-102-92-84-84-91
-108-107-107-107-107-96-87-86-85-85-9595-95
.75-81-104-103-94-123-99-86-88-89-98-109

المطلوب:

تمثيل البيانات أعلاه بفكرة الاغصان والأوراق، فمن الممكن عمل الغصن على اعتباره يعبر عن رقم العشرات والمئات، والأوراق يعبر عنها رقم الاحاد في قيمة كل مشاهدة (نتيجة كل طالب) .

فالسطر الأول يحتوي على النتائج التي جاءت اقل من 80.

والسطر الثاني يحتوي على النتائج التي جاءت تتراوح بين اقل من 80 واقل من90 .

والسطر الثالث يحتوي على النتائج التي جاءت تتراوح بين 90 واقل من 100 وهكذا...

يجب التذكير بان الأوراق التي يحددها رقم الاحاد في قيمة كل مشاهدة، سوف يتفاوت عددها من غصن الى اخر.

7	**85**
8	**23 44 55 67 98 61**
9	**50 12 33 55 55 56 68 94**
10	**00 22 66 77 77 89 34**
11	**02 34 79**
12	**53**

الان يجب ترتيب الأوراق من الأصغر الى الأكبر كما يلي:

7	**58**
8	**12344557986**
9	**01233455555689**
10	**00223466777789**
11	**023479**
12	**35**

يمكن من خلال الاغصان والأوراق معرفة ان:

- المشاهدة الصغرى 85 وهو ما يظهر في السطر الأول حيث الغصن 7 والورقة 5
- المشاهدة الكبرى 125 وهو ما يظهر في الطر الأخير حيث الغصن 12 والورقة 5.
- عدد كبير من المشاهدات جاء في الغصن0 والغصن 9.
- توزيع النتاج تقريبا متماثل او اقرب الى الاعتدالية وتتحدد وحدات الاغصان بالأوراق بناء على طبيعة وحجم البيانات التي جمعت، فليس بالضرورة ان تكون فئات العشرات او المئات هي الغصن.

مقاييس النزعة المركزية– الوسط الحسابي

محتوى المحاضرة:

- **مفاهيم أساسية حول مقاييس النزعة المركزية.**
- **الوسط الحسابي.**
- **تعريف الوسط الحسابي.**
- **حالة البيانات غير المبوبة.**
- **حالة البيانات المبوبة.**
- **الطريقة المباشرة**
- **الطريقة غير المباشرة.**
- **الوسط الحسابي المرجح.**
- **مزايا وعيوب الوسط الحسابي.**

مقاييس النزعة المركزية– الوسط الحسابي

تسمى بمقاييس النزعة المركزية لان الدرجات او البيانات تنوع الى الاقتراب الى المركز او تبتعد عنه، لو افترضنا ان البيانات مجمعة في دائرة ومركز هذه الدائرة هو مثلا المتوسط فان هذه البيانات قد تقترب الى المركز أي تنزع اليه وقد تبتعد عنه، ولذلك فإن المتوسط يتأثر بالقيم المتطرفة أي المرتفعة او المنخفضة عنه بكثير.

وباعتبار ان هذه البيانات تميل في الغالب الى المركز فانه يمكن استخدام هذه "القيمة المركزية" لتمثيل هذه المجموعة من البيانات والمقاييس المستخدمة في التعرف على هذه القيمة المركزية والتي تسمى مقاييس النزعة المركزية. يجب توفر الشروط التالية لكي يكون المقياس جيدا:

- ان يعتمد المقياس في حسابه على كل المشاهدات.
- ان يكون المقياس سهل الحساب والفهم.
- ان يتوفر فيه القابلية للتعامل الجبري.
- ان يتأثر بالقيم الشاذة او المتطرفة.

من بين هذه المقاييس وأكثرها شيوعا، الوسط الحسابي، الوسيط، والمنوال. ولا تتوفر كل الصفات السابق ذكرها في مقياس واحد، ولكن كل مقياس من هذه المقاييس يفضل استخدامه في حالات معينة ولا يفضل استخدامه في حالات أخرى.

أولا– الوسط الحسابي: يعرف الوسط الحسابي لمجموعة من القيم بانه مجموع قيم الظاهرة مقسوما على عدد مفرداتها، ويرمز له بالرمز $\bar{X}$

أ- **حالة البيانات غير المبوبة:**

ويحسب وفق القانون التالي:

$$\bar{X}=\frac{\sum x}{n}$$

- حيث x يمثل قيم الظاهرة
- حيث n تمثل عدد المفردات.
- حيث $\bar{X}$ يمثل الوسط الحسابي للظاهرة X .

مثال: احسب الوسط الحسابي لعلامات 10 طلاب في مقياس ما والذين كانت بيناتهم كالتالي:

14 14 10 15 12 14 16 15 12 13

الحل:

$$125 = 14+14\ 10+\ 15+\ 12+\ 14\ 16+\ 15\ +12+\ 13$$

$$12.5 =\frac{125}{10} =\bar{X}$$

ب- **حالة البيانات المبوبة:** يحسب وفق طريقتين:

- **الطريقة المباشرة:** يتخذ مركز الفئة كقيمة تضرب في التكرار المقابل، ثم تجمع القيم المحصل عليها وتقسم على مجموع التكرارات وتعطى وفق الصيغة:

$$\bar{X} = \frac{x1n1+x2n2+\cdots xnnn}{N} = \frac{\sum nixi}{N}$$

مثال: يعرض الجدول التالي توزيع اوزان مجموعة من الرياضيين:

فئات الوزن	60-58	62-60	64-62	66-64	68-66	70-68	72-70
عدد الرياضيين	06	08	10	12	09	14	06

المطلوب: اوجد الوسط الحسابي؟

الحل:

لحساب الوسط الحسابي يتم اتباع الخطوات التالية:

- إيجاد مجموع التكرارات.
- حاب مراكز الفئات.
- ضرب مركز كل فئة في التكرار المناظر لها.
- حساب مجموع التكرارات في مراكز الفئات.

فئات الوزن	التكرارات i	مراكز الفئات x_i	$n_i x_i$
60-58	06	88=2/60+58	352
62-60	08	61	488
64-*62*	10	63	630
66-64	12	65	780
68-66	09	67	603
70-68	14	69	966
72-70	06	71	426
المجموع	65	/	4245

$$\bar{X} = \frac{\sum nixi}{N} = \frac{4245}{65} = 65.30 \text{ kg}$$

- **الطريقة غير المباشرة:** يتم اتباع الخطوات التالية
- تحديد مراكز الفئات.
- تحديد متوط فرضي ويفض ان يكون مركز الفئة ذات اكبر تكرار.
- حساب انحراف مراكز الفئات عن هذا الوسط الفرضي.
- ضرب كل انحراف بالفئة التي تقابله
- تستخدم العلاقة التالية:

$$\bar{X} = B + \frac{\sum nidi}{N}$$

حيث di هو انحراف مركز الفئة عن الوط الفرضي بمعنى $di = x_i - B$

مثال: نفس معطيات المثال السابق

- ليكن $B=69$

فئات الوزن	التكرارات n_i	مراكز الفئات x_i	$d_i = (x_i - B)$	$n_i d_i$
60-58	06	88	19	114
62-60	08	61	-8	-64
64-62	10	63	-6	-60
66-64	12	65	-4	-48
68-66	09	67	-2	-18
70-68	14	69	0	0
72-70	06	71	2	12
المجموع	65	/	/	64

بالتعويض في العلاقة:

$$\bar{X} = B + \frac{\sum nidi}{N} = 69 + \frac{64}{65}$$

$$= 69 + 0.98$$

$$= 69.98\ kg$$

- **الوسط الحسابي المرجح:**

في بعض الأحيان يكون لكل قيمة من قيم المتغير أهمية نسبية تسمى ترجيحات، إن عدم أخذ هذه الترجيحات في الاعتبار عند حساب الوسط الحسابي، يؤدي الى الحصول على قيمة معبرة عن الوسط الحسابي بشكل غير دقيق.

مثلا لو اخذنا 12 طالبا وسجلنا علاماتهم في امتحان مقياس المنهجية مثلا، وسجلنا أيا عدد ساعات المراجعة في الأسبوع.

الطالب	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***
العلامة *x*	***17***	***18***	***16***	***15***	***16***	***12***	***13***	***14***	***14***	***16***	***12***	***15***
عدد ساعات المراجعة *w*	***20***	***21***	**19**	***19***	**19**	**17**	***18***	**19**	**18**	**18**	**17**	**17**

- الوسط الحسابي غير المرجح للعلامة الحاصل عليها الطالب هي:

$$\bar{X} = \frac{\sum x}{n} = \frac{17+18+16+1516+12+13+1414+16+12+15}{12} = 14.83$$

- وإذا اردنا حساب الوسط الحسابي للعلامات x المرجحة بعدد ساعات المراجعة w يتم تطبيق المعادلة التالية:

$$\bar{W} = \frac{\sum xw}{\sum w} = \frac{17\times20+18\times21+16\times19+\cdots\ldots15\times19}{20+21+\cdots\ldots+17}$$

$$= \frac{3314}{222}$$

$$= 14.94$$

وهو الوسط الحسابي المرجح والذي يظهر بأكثر دقة من الوسط الحسابي غير المرجح.

- **مزايا وعيوب الوسط الحسابي:**
- يعتبر الوسط الحسابي أبسط مقاييس النزعة المركزية حسابا وأكثرها استخداما.
- يأخذ المتوسط الحسابي بين الاعتبار جميع قيم الظاهرة المدروسة.
- مجموع انحرافات القيم عن متوسطها الحسابي يساوي0

$$\sum_{i=1}^{n}(xi - \bar{X}) = 0$$

- يتأثر بالقيم الشاذة والمتطرفة.
- يصعب حسابه في حالة البيانات الوصفية.
- يصعب حسابه في حالة الجداول التكرارية المفتوحة.

مقاييس النزعة المركزية– الوسيط

محتوى المحاضرة

- تعريف الوسيط.
- طرق حساب الوسيط.
- حساب الوسيط للبيانات المبوبة.
  - في حالة القيم الفردية.
  - في حالة القيم الزوجية.
- حساب الوسيط للبيانات المبوبة.
- مميزات الوسيط.

مقاييس النزعة المركزية– الوسيط

ثانيا– الوسيط:

1- **تعريف الوسيط:** يفضل عل الكثير من المقاييس الإحصائية لعدم تأثر قيمته بالمفردات المتطرفة في التوزيع، هو القيمة التي تقسم مجموعة القيم بعد ترتيبها تصاعديا او تنازليا الى قسمين متساويين، بحيث يكون عدد القيم الأصغر منها مساويا لعدد القيم الأكبر منها، ويرمز له بالرمز (m).

2- **طرق حساب الوسيط:**

2–1–**حساب الوسيط للبيانات غير المبوبة:** يستخرج الوسيط من البيانات غير المبوبة، حسب عدد القيم المدروسة.

أ– **في حالة قيم فردية:** اذا كان عدد القيم فرديا فان الوسيط هو القيمة التي تشغل الرتبة وفق العلاقة $\frac{n+1}{2}$

مثال: لتكن لدينا القيم التالية:

90 50 70 40 100 60 80

المطلوب: أوجد قيمة الوسيط؟

الحل:

– لحساب قيمة الوسيط نرتب القيم ترتيبا تصاعديا فتصبح:

40 50 60 70 80 90 100

– نستخرج رتبة الوسيط :=$\frac{n+1}{2}$

=4$\frac{7+1}{2}$

- نلاحظ أن الوسيط يشغل الرتبة الرابعة ومنه قيمة الوسيط : m= 70

ب- **في حالة قيم زوجية:** اذا كان عدد القيم زوجيا فان الوسيط يكون أي قيمة بين القيمتين اللتان تقعان في الوسط، وتشغلان المركز $\frac{n}{2}$

والمركز $\frac{n+1}{2}$

مثال: لتكن لدينا القيم التالية:

30 60 40 70 50 80

لحساب قيمة الوسيط نرتب القيم تصاعديا فتصبح

30 40 50 60 70 80

نحدد موضع الوسيط من خلال المشاهدة يتضح أنه بين القيمة الثالثة والرابعة ومنه:

$$m= \frac{50+60}{2} = 55$$

2-2-**حالة البيانات المبوبة:** يتم إيجاد قيمة الوسيط بالحساب وبالرسم ولا يهم ان كانت اطوال الفئات متساوية او مختلفة، ويتم ذلك باتباع الخطوات التالية:

- نكون جدول تكراري متجمع صاعد او نازل.
- إيجاد موقع الوسيط وفق العلاقة $C_1=\frac{\sum f}{2}$
- إيجاد قيمة الوسيط من العلاقة الاتية: $m=L+\frac{c1-c2}{c3}.h$
- *حيث:*
- *M : رمز الوسيط.*
- *L : بداية الحد الأدنى لفئة الوسيط*
- C_1 : ترتيب الوسيط.

- C_2 : تكرار متجمع صاعد للفئة الوسيطية.
- C_3 : التكرار الأصلي لفئة الوسيط
- h : طول فئة الوسيط.
- يستخدم هذ القانون في حالة الجدول الصاعد.
- يلاحظ ان قيمة الوسيط لا بد وأن تقع داخل حدود فئة الوسيط، أي لا تل عن بداية فئة الوسيط ولا تزيد عن نهايتها.

مثال: من خلال البيانات الموضحة في الجول التالي اوجد قيمة الوسيط لتوزيع 20 عداء حسب التوقيت المقطوع في سباق ما بالثواني.

فئات الزمن c	عدد العدائينf	ك.م.ص	اقل من الحد الاعلى للفئة
10-20	03	3	أقل من 20
20-30	6	9	أقل من 30
30-40	10	19	أقل من 40
40-50	15	34	أقل من 50
50-60	8	42	أقل من 60
60-70	5	47	أقل من 70
70-80	3	50	أقل من 80
Σ	50	-	-

- بعد تحويل الجدول الى جدول تكراري متجمع صاعد
- $C_1=\frac{\sum f}{2} = \frac{50}{02} = 25$ وفق العلاقة نجد ان ترتيب الوسيط 25

- يتم تحديد الفئة الوسيطية ، أي ان نبحث في عمود التكرار التجميعي الصاعد عن القيمة 25 او اكبر منها مباشرة.
- يلاحظ من عمود التكرار التجميعي الصاعد ان القيمة 25 غير موجودة وبالتالي الأكبر منها مباشرة هي 34 وتقابلها الفئة 50−40
- حساب قيمة الوسيط:

$$m=L+\frac{c1-c2}{c3}.h= 40+\frac{25-19}{15}\times 10$$
$$=44$$

- يمكن إيجاد قيمة الوسيط بالرسم باتباع الخطوات التالية:
- يرسم منحنى متجمع صاعد او نازل.
- تحدد ترتيب الوسيط على المحور العمودي.
- نوجد قيمة الوسيط على المحور الافقي.

3- **مميزات الوسيط: ي**تصف الوسيط بعدة خصائص أهمها:

- لا يتأثر بالقيم المتطرفة وبالتالي فانه يعتبر أصلح المقاييس عند وجود مثل هذه القيم.
- يمكن إيجاد الوسيط من الرسم.
- يمكن استخدامه في حالة البيانات النوعية.
- يمكن حسابه من الجداول التكرارية المفتوحة.

مقاييس النزعة المركزية– المنوال

محتوى المحاضرة

- تعريف المنوال.
- طرق حساب المنوال.
- حساب المنوال من البيانات غير المبوبة.
- حساب المنوال من البيانات المبوبة.
- مميزات المنوال.

مقاييس النزعة المركزية– المنوال

ثالثا المنوال:

1–**تعريف المنوال:** هو القيمة التي تتكرر أكثر من غيرها، ويرمز له بالحرف D.

2–**طرق حساب المنوال:**

2–1– حساب المنوال من البيانات غير المبوبة:

في حالة تكرار رقم واحد يتم اختياره كمنوال، اما في حال تكرار رقمين بنفس عدد مرات التكرار يتم اختيارهما معا كمنوال، اما إذا زاد أحدهما عن الآخر يتم اختيار ذو التكرار الأكبر، وفي حال عدم تكرار أي رقم فانه لا يوجد منوال.

مثال: أحسب المنوال في كل من الحالات التالية:

7 8 9 8 10 8 12 – D = 8

10 12 10 15 12 10 --- D = 10

15 16 15 20 16 30------- D = 15، 16

20 30 40 140 50 60 -----لا يوجد

2–2– حساب المنوال من البيانات المبوبة:

يتم ايجاد قيمة المنوال بالحساب وبالرسم

أ– **بالحساب:** نستخدم العلاقة التالية:

$$D=L+\frac{\Delta 1}{\Delta 1+\Delta 2}.h$$

حيث أن:

- L: الحد الأدنى لفئة المنوال.
- Δ_1 : الفرق بين اكبر تكرار والسابق له.

- Δ_2 : الفرق بين اكبر تكرار واللاحق له.
- h : طول فئة المنوال.
- تعرف فئة المنوال بانها الفئة الأكبر تكرار

مثال: أوجد المنوال من الجدول التالي:

الفئات	10−	20−	30−	40−	50−	60−	80−70
التكرارات	5	12	22	38	22	12	5

الحل:

الفئة	التكرار
10−	5
20−	12
30−	22
40−	38
50−	22
60−	12
80−70	5

- نحدد الفئة المنوالية من خلال أكبر رقم في عمود التكرار.
- نحدد الحد الأدنى للفئة المنوالية وهو بدايتها وهو L=40.
- نحدد تكرار الفئة المنوالية، وتكرار الفئة التي تسبق الفئة المنوالية، وتكرار الفئة التي تلي الفئة المنوالية.

- نحدد طول الفئة

بالتعويض في العلاقة:

$$D=40+\frac{16}{16+16}\times 10 = 45$$

ب- **بالرسم:** يتم إيجاد قيمة المنوال بالرسم من المدرج التكراري، ويمكن الاكتفاء برسم ثلاث مستطيلات، مستطيل يمثل فئة المنوال- الفئة التي تقابل أكر تكرار- ومستطيل يمل الفئة السابقة وآخر يمل الفئة اللاحقة لها، ثم نصل رؤوس المستطيلات ببعضها فتتقابل في نقطة نُسقط منها عمود على المحور الافقي، فتكون هي قيمة المنوال.

3-مميزات المنوال:

- لا يأخذ بالاعتبار جميع البيانات المعطاة وبالتالي فهو لا يتأثر بالقيم المتطرفة.
- يمكن حسابه بيانيا.
- يمكن ان يوجد اكر من منوال لتوزيع واحد.
- يمكن حسابه من الجداول الإحصائية المفتوحة.
- يعتبر أفضل المتوسطات لوصف الظواهر النوعية.

مقاييس التشتت

محتوى المحاضرات

- مدخل مفاهيمي لمقاييس التشتت.
- مقاييس التشتت –المدى
- أمثلة تطبيقية لحساب المدى.
- مقاييس التشتت– الانحراف المتوسط.
- أمثلة تطبيقية لحساب الانحراف المتوسط.
- خصائص الانحراف المتوسط.
- مقاييس التشتت– التباين.
- تباين البيانات غير المبوبة.
- تباين البيانات المبوبة.
- أمثلة تطبيقية لحساب التباين.
- مقاييس التشتت– الانحراف المعياري
- الانحراف المعياري للبيانات غير المبوبة
- الانحراف المعياري للبيانات المبوبة.
- امثلة تطبيقية لحساب الانحراف المعياري.
- مقاييس التشتت– معامل الاختلاف.

مقاييس التشتت

تسمح لنا مقاييس النزعة المركزية من متوسط ووسيط ومنوال بالحصول على القيم المتوسطة للبيانات او على تجمعها، غير ان هذه المقاييس لا تكفي لوحدها لمعرفة الصفات الإحصائية اللازمة لوصف الظواهر، لان الفروق بين قيم الظواهر قد تزدادا أو تنقص رغم تساوي المتوسطات لهذه الظواهر. لتوضيح ذلك نفترض ان قسمين دراسيين، يحتوي كل قسم على 05 طلبة، وكانت نتائج الطلبة كالتالي:" (راتول، 2018، ص137).

القسم الأول	**10**	**12**	**13**	**10**	**10**
القسم الثني	**3**	**20**	**20**	**5**	**7**

ان المتوسطات الحسابية لنتائج القسمين تظهر متساوية 11، يعني هذا ان المستوى البيداغوجي للقسمين متساو، غير انه لو لاحظنا علامة كل طالب لوجدنا ان كل طلبة القسم الأول ناجحون بينما لم ينجح في القسم الثاني سوى طالبين.

فالحقيقة اذن ان مستوى القسمين غير متساو، وهو ما يثبت ان مقاييس النزعة المركزية لا تعطي فكرة وافية عن اختلاف قيم الظواهر، ولا تحقق كل الأغراض التي نرغب الوصول اليها من دراستنا، فلو لاحظنا الفرق بين اكبر علامة واصغر علامة في القسم الأول لوجدناها = 03، بينما يقدر الفرق بين اكبر علامة واصغر علامة في القسم الثاني = 17، أما اكبر انحراف عن الوسط الحابي في القسم الأول فهو = 02، واكبر انحراف عن الوسط الحسابي في القسم الثاني = 09، وهو ما يعني ان نتائج القم الثاني اكبر تباعدا أي تشتتا عن بعها

وعن وسطها الحسابي، على عكس نتائج القسم الأول فإنها اقل تشتتا، من هنا جاءت أهمية مقاييس التشتت.

تعريف: إذا كانت بيانات ظاهرة ما متباعدة وغير متجانسة فنقول ان مفرداتها مشتتة وغير مركزة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمتها قريبة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة. أي ان التشتت هو مدى تباعد مجموعة من القيم عن بعضها البعض او عن القيمة التي تمثل مركز تلك المجموعة. كلما ارتفعت قيم مقاييس التشتت دل ذلك على درجة كبيرة من التباعد والاختلاف بين قيم البيانات، وكلما كانت صغيرة دل ذلك على أن الاختلاف بين قيم البيانات قليل. يقاس تشتت البيانات بعدة مقاييس منها:

1- **المدى:** يعرف المدى بأنه الفرق بين اكبر قيمة وأصغر قيمة لمجموعة من البيانات سواء كانت قيم هذه المجموعة مبوبة أو غير مبوبة ويرمز له بالرمز R .

مثال(01): من المثال السابق أوجد المدى؟

الحل: نحسب المدى للقسمين:

- القسم الأول: $R= X_{max} - X_{min} = 13-10=03$
- القسم الثاني: $R= X_{max} - X_{min} = 20-03=17$

إن بيانات القسم الأول اقل تشتتا وبالتالي فان وسطها الحسابي يمثلها تمثيلا مقبولا، بينما بيانات القسم الثاني متباعدة أي أكثر تشتتا وبالتالي فان وسطها الحسابي لا يمثلها تمثيلا جيدا.

مثال (02): أوجد المدى للبيانات التالية:

8 21 22 08- 18 24 76

الحل:

R= 76−(−8)=84

نلاحظ ان المدى هنا قد تأثر بكل كبير جدا بالقيم المتطرفة، باعتبار ان معظم البيانات متقاربة باستثناء القيمة 76 والقيمة −8، فاذا استبعدنا هذه القيم المتطرفة فان المدى يصبح:

R= 24−8=16وبسبب هذا العيب فان المدى كمقياس للتشتت يعتبر مقياسا مضللا ، لأنه يتوقف على اكبر قيمة واصغر قيمة ، وبالتالي فهو لا يعكس بكل فعلي تشتت البيانات.

2−الانحراف المتوسط:

يقصد به المتوسط الحسابي للقيم المطلقة لانحرافات القيم عن متوسطها الحسابي، ويرمز له بالرمز E_x .

مثال (01): أوجد انحرافات المتوسط للقيم التالية؟

10 6 14 8 07

الحل:

- الوسط الحسابي: $\bar{x}=\frac{\sum x}{n}$ = 9
- حساب الانحرافات:

- $\sum|xi-\bar{x}|$= 1+3+5+1+2=12

$Ex=\frac{12}{05}=2.4$

أما إذا كانت البيانات **مبوبة** في جداول توزيع تكرار فان الانحراف المتوسط لها يعطى وفق العلاقة التالية:

$$E_{x=}\frac{\sum_{i=1}^{n} fi|XI-\bar{X}|}{\sum_{i=1}^{n} fi}$$

مثال (02): اوجد الانحراف المتوسط للبيانات التالية :

7 15 6 12 10 8 18 4

الحل:

$$\bar{X}=\frac{\sum xi}{n}=\frac{80}{8}=10$$

$\|Xi-\bar{X}\|$	X_i
6	4
8	18
2	8
0	10
2	12
4	6
5	15
3	7
28	المجموع

- حساب الانحراف:

$$E_x = \frac{\sum|Xi-\bar{X}|}{n} = \frac{28}{5} = 5.6$$

مثال (03): أوجد تشتت البيانات المبوبة في جدول التوزيع التكراري باستخدام الانحراف المتوسط؟

الفئة	2-0	4-2	6-4	8-6	المجموع
التكرار	2	3	4	3	12

الحل:

الفئة	التكرار f_i	X_i	$X_i.f_i$	$\|Xi-\bar{X}\|$	$fi\|Xi-\bar{X}\|$
2-0	2	1	2	3.33	6.66
4-2	3	3	9	1.33	4
6-4	4	5	20	0.67	2.68
8-6	3	7	21	2.67	8.01
المجموع	12		52		21.35

$$\bar{X} = \frac{\sum fi\,Xi}{\sum fi} = \frac{52}{12} = 4.33$$

$$E_x = \frac{\sum_{i=1}^{n} fi|XI-\bar{X}|}{\sum_{i=1}^{n} fi} = \frac{21.35}{12} = 21.78$$

خصائص الانحراف المتوسط:

- يعتمد في حسابه على جميع القيم ولي القسمة الكبرى والصغرى فقط.
- لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة.
- بتأثر بالقيم المتطرفة، لان انحرافها عن المتوسط الحسابي يكون كبيرا.

3- التباين:

لتجنب عيب الانحراف المتوسط وهو عدم الخضوع للعمليات الجبرية- بسبب القيمة المطلقة- من خلال حساب الانحراف المتوسط كمتوسط مجموع الانحرافات مع اهمال الإشارة السالبة. فإنه يتم إيجاد متوسط مربعات انحرافات القيم- الفروق- بين قيم المتغير الاحصائي ومتوسطها الحسابي، ويسمى هذا المتوسط بالتباين.

أ- **تباين البيانات غير المبوبة:** اذا كان لدينا البيانات التالية:

$X_1, X_2, X_3, \ldots\ldots\ldots X_n$ فإن التباين لهذه البينات يعطى بالعلاقة التالية:

$$V_x = \frac{\sum_{i=1}^{n}(Xi-\overline{X})^2}{n}$$

في بعض الأحيان عندما يكون المتوسط الحسابي للبيانات عبارة عن كسر فإن حساب التباين يكون متعبا وعرضة للأخطاء الحسابية لذلك يمكن كتابة العلاقة السابقة على النحو التالي:

$$V_x = \frac{\sum Xi^2}{n} - \overline{X}^2$$

مثال: هذه علامات لعينة من الطلبة في امتحان مقياس الإحصاء:

10 15 12 13 16 11 12 7

المطلوب: أوجد التباين؟

الحل:

- حساب المتوسط الحسابي:

$$\overline{X}=\frac{\sum xi}{n}=\frac{96}{8}=12$$

$\|Xi - \bar{X}\|^2$	$\|Xi - \bar{X}\|$	X_i	الطلبة
25	−5	7	1
0	0	12	2
1	−1	11	3
16	4	16	4
1	1	13	5
0	0	12	6
9	3	15	7
4	−2	10	8
56		96	المجموع

- حساب التباين:

$$V_x = \frac{\sum(Xi-\bar{X})^2}{n} = \frac{56}{8} = 7$$

ب- **تباين البيانات المبوبة:** في حال البيانات المبوبة فإن العلاقة تصبح على الكل التالي:

$$V_x = \frac{\sum fi.Xi^2}{\sum fi} - \bar{X}^2$$

مثال: فيما يلي توبع علامات 18 طالبا.

المطلوب: أحسب التباين؟

الفئات	f_i	Xi	X_i^2	$f_i X_i^2$	fiX_i
16−12	2	14	196	392	28
20−16	3	18	324	972	54
24−20	5	22	484	2420	110
28−24	6	26	676	4056	156
32−28	4	30	900	3600	120
المجموع	20	/		11440	468

$\bar{X} = \frac{\sum fi.Xi}{n} = \frac{468}{20} = 23.4$

$V_x = \frac{\sum fi.Xi^2}{\sum fi} - \bar{X}^2$

$= \frac{11440}{20} - 547.56$

$= 572 - 547.56$

$= 24.44$

4-الانحراف المعياري:

من اهم المقاييس الإحصائية للتشتت، وأكثرها استعمالا كونه يعطي فكرة ليمة ومنطقية عن ظاهرة التشتت، يقيس مدى تباعد أو اختلاف أو انحراف قيم المتغير الاحصائي عن وسطه الحسابي.$\bar{X}^2$

يعرف بأنه الجذر التربيعي لمتوسط مجموع مربع انحراف القيم عن متوسطها، أي أنه الجذر التربيعي للتباين.

أ- الانحراف المعياري للبيانات غير المبوبة: يتم حسابه وفق العلاقة:

$$S_x = \sqrt{\frac{\sum(Xi-\bar{X})^2}{n}}$$

أما الصيغة المختصرة للانحراف المعياري للبيانات غير المبوبة فهي:

$$S_x = \sqrt{\frac{\sum Xi^2}{n} - \bar{X}^2}$$

مثال(01): أوجد الانحراف المعياري للقيم التالية:

13 10 23 12 21 11

الحل:

- حساب المتوسط الحسابي:

$$\bar{X} = \frac{\sum xi}{n} = 15$$

- تربيع القيم: 121 441 144 529 100 169
- حساب الانحراف المعياري:

$$S_x = \sqrt{\frac{\sum Xi^2}{n} - \bar{X}^2}$$

$$S_x = \sqrt{\frac{1504}{6} - 225}$$

$$S_x = 5.06$$

ب- **الانحراف المعياري للبيانات المبوبة:** يتم حسابه وفق العلاقة التالية:

$$S_x = \sqrt{\frac{\sum fi\,(Xi-\bar{X})^2)}{\sum fi}}$$

وبالصيغة المختصرة يكتب بهذا الكل:

$$S_x = \sqrt{\frac{\sum fiXi^2}{\sum fi} - \bar{X}^2}$$

مثال(02): اوجد الانحراف المعياري للبيانات المبوبة في الجدول الاحصائي التالي:

الفئة	20-10	30-20	40-30	50-40	60-50	70-60	المجموع
التكرار	6	5	8	7	4	2	32

الحل:

الفئة	n_i	x_i	$n_i x_i$	x_i^2	$n_i x_i^2$
20-10	6	15	90	225	8100
30-20	5	25	125	625	15626
40-30	8	35	280	1225	78400
50-40	7	45	315	2025	99225
60-50	4	55	220	3025	48400
70-60	2	65	130	4225	16900
المجموع	32	/	1160	11350	266651

- حساب المتوسط الحسابي:

$$\bar{X} = \frac{\sum niXi}{N} = \frac{1160}{32} = 36.25$$

- حساب الانحراف المعياري:

$$S_x = \sqrt{\frac{\sum fiXi^2}{\sum fi} - \bar{X}^2}$$

$$S_x = \sqrt{\frac{266651}{32} - 1314.06}$$

$$= 83.77$$

5-معامل الاختلاف:

من خلال ما سبق ذكره يتضح أنه إذا استخدمنا هذ المقياس- الانحراف المعياري- فإن المقارنة بين تشتت ظاهرتين أو أكثر تكون واقعية وواقعية فقط إذا كانت الظواهر من نوعية واحدة ولها متوسطات متساوية، أي يمكن مقارنة تشتت درجات مادة ما بدرجات مادة أخرى او مقارنة تشتت دخل مجموعة من العمال بدخل مجموعة أخرى وتكون المقارنة أكثر واقعية إذا كانت المتوسطات متساوية أو قريبة من بعضها.

أما إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها متباعدة، فان المقارنة اعتمادا على الانحراف المعياري ستكون غير منطقية وغير واقعية، ولهذا السبب وجدت مقاييس أخرى سميت مقاييس التشتت النسبي، تعتمد على تمييز البيانات وتقيس التشتت كنسبة مئوية للمتوسط، أهم هذه المقاييس هو معامل الاختلاف.

$$\textbf{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{المتوسط الحسابي}} \times 100.$$

الارتباط والانحدار الخطي البسيط

فكرة عن الارتباط بين المتغيرات:

إن اقتران التغير في ظاهرة بالتغير في ظاهرة أخرى يعرف بالارتباط، ويهتم بموضوع الارتباط بقياس العلاقة بين الظواهر المختلفة، ظاهرتين او أكثر وذلك لمعرفة ما اذا كان تغير احدها او مجموعة منها يقترن بتغبر الأخرى.

مثال ذلك الارتباط بين مساحة الدائرة ونصف قطرها، نلاحظ هنا ان زيادة قيمة نصف القطر للدائرة يقترن به زيادة مساحة الدائرة، ونقول هنا ان العلاقة او الارتباط طردي بين المتغيرين أي بين نصف القطر والمساحة.

ومن الأمثلة أيضا الارتباط بين الذكاء والتحصيل الدراسي لدى التلاميذ ففي الغالب تجد ان أصحاب المستوى المرتفع في الذكاء أكثرهم لهم تحصيل مرتفع وكذا أصحاب المستوى المنخفض في الذكاء أكثرهم لهم تحصيل اقل. ويقال هنا أيا ان العلاقة طردية بين المتغيرين، ونقصد بين الذكاء والتحصيل الدراسي، وهنا تكون الزيادة في المتغير الأول مقترنة بالزيادة في المتغير الثاني، ونقول ان التغير الاقتراني موجبا او طرديا.

واحيانا تأتي العلاقة عكسية بين المتغيرين، مثل علاقة التوافق النفسي بالقلق ونقصان حجم قطعة الثلج تبعا لزيادة درجة الحرارة، وهنا تكون الزيادة في المتغير الأول مقترنة بالنقص في المتغير الثاني او العك، ونقول ان التغير الاقتراني سلبيا او عكسيا.

والتغير الاقتراني سواء كان عكسيا او طرديا أي سالبا او موجبا يقاس بقوانين لمعامل الاقتران او الارتباط . وتحسب قيمة معامل الارتباط باستخدام قوانين مختلفة مثل معامل الارتباط

بيرسون، للدرجات الخام، ومعامل الارتباط سبيرمان للرتب، وقيمة معامل الارتباط المحسوبة باي من هذه الطرق تتراوح بين+1، −1.

اذا وصلت قيمة معامل الارتباط الى +1 قلنا ان الارتباط تام وموجب بين المتغيرين، واذا وصلت قيمة معامل الارتباط بين المتغيرين الى −1 قلنا ان الارتباط تام سالب بين المتغيرين.

وعلى أي حال فان اقتراب قيمة معامل الارتباط من الواحد (موجبا كان او سالبا) تشير الى قوة الارتباط بين المتغيرين، كما ان اقتراب معامل الارتباط من الصفر يدل على عف الارتباط بين المتغيرين.

الارتباط بشكل عام مصطلح يشير بمعناه الحرفي الى المتشابهات في ظاهرة من الظواهر، او درجة التلازم في التغير بين متغير او أكثر وبالتالي فانه يستخدم لقياس العلاقة بين مجموعتين او أكثر من البيانات.

الارتباط البسيط:

هو ذلك الارتباط الذي يدرس العلاقة بين متغيرين اثنين فقط، مثل الطول كمتغير والوزن كمتغير ثان، والعمر الزمني كمتغير والقوة العضلية كمتغير الى ذلك من المتغيرات، ويطلق على الارتباط البسيط في بعض الأحيان اسم الارتباط ذو المتغيرين حيث يعتبر احد المتغيرين تابعا والأخر مستقلا.

الارتباط المتعدد:

هو ذلك الارتباط الذي يدر العلاقة بين أكثر من متغيرين، احدهما متغير غير مستقل والمتغيرات الأخرى مستقلة كان ندرس مثلا العلاقة بين التحصيل الدراسي كمتغير غير مستقل وكل من الذكاء وعدد ساعات المذاكرة كمتغيرين مستقلين.

الارتباط الجزئي:

هو ذلك الارتباط الذي يدرس العلاقة بين المتغير التابع ومتغير مستقل معين مع عزل تأثير جميع المتغيرات المستقلة الأخرى، كان ندرس مثلا العلاقة بين التحصيل الدراسي كمتغير غير مستقل والمستوى الاقتصادي والاجتماعي كمتغير مستقل، مع عزل تأثير الذكاء كمتغير مستقل أيضا.

الارتباط الخطي:

وهو يعني ان ما يحدث من تغير في أحد المتغيرات يتبعه تغير بمقدار ثابت بالنسبة للمتغير الاخر، بمعنى ان الزيادة في احد المتغيرات يصاحبها زيادة في المتغير الاخر او ان الزيادة في احد المتغيرات يصاحبها نقص في المتغير الاخر.

الارتباط الإيجابي:

تطرقنا اليه سلفا وهو ارتباط يؤدي فيه تزايد معين الى تزايد في المتغير الاخر وفي حالة هذا النوع من الارتباط تكون لمعامل الارتباط قيما موجبة تتراوح من +0.01 الى +1 كما ذكرنا.

الارتباط السلبي:

هو ارتباط يؤدي فيه تزايد متغير معين، الى تناقص في المتغير الاخر، وفي حالة هذ النوع من الارتباط تكون لمعامل الارتباط قيم سالبة تتراوح من −0.01 الى −1 .

معامل الارتباط:

عبارة عن مؤشر عددي يستخدم للتعبير الكمي عن العلاقة الممتدة بين متغيرين او أكثر، هذ المؤشر يسمى بمعامل الارتباط، حيث يرمز لهذا المعامل في إحصاء العينة بالرمز r وفي إحصاء المجتمع الأصلي بالرمز **r**.

منطق معامل الارتباط:

يستخدم معامل الارتباط لتقدير العلاقة بين متغيرين او أكثر تقديرا كميا، ولا يتضمن معانل الارتباط المحسوب بين متغيرين أي صورة لوجود سببية بين المتغيرين، بمعنى انه عندما نقوم بحاب معامل الارتباط بين المتغيرين أ و ب ، فإننا لا نستطيع ان نقرر او نستنتج ان أ هو البب و ب هو النتيجة او ان نؤكد ان ب متغير مستقل و أ متغير تابع فقد يكون المتغيران أ و ب متغيرين تابعين وان هناك متغير ثالث او رابع او عددا من المتغيرات تؤثر كمتغيرات مستقلة في العلاقة بين أ و ب. لذا لا تتضمن العلاقة الارتباطية بين أي متغيرين أكثر من وجود قدر محسوب في شكل معامل احصائي يكشف عن مقدار التلازم في التغير بين متغيرين أدى لوجود تباين مشترك بينهما، هذا التباين المشترك يخلو تماما من اية معلومات عن السبب والنتيجة (العليّة).

من الواجب في الدراسات الارتباطية ان نهتم بتفسير أسباب العلاقة بين المتغيرات وان نبتعد تماما عن التفسيرات التي تتناول تأثير متغير على اخر او تأثير متغير او متغيرات على ظاهرة معينة، فالمنهج الارتباطي يخلو تماما من اية دلالات تتعلق بالسبب والنتيجة العلية التي تدخل في نطاق البحوث التجريبية.

معامل الارتباط بيرسون:

ويعرف باسم معامل الارتباط بطريقة العزوم، او معامل الارتباط التتابعي بطريقة بيرسون(PPM) ويستخدم هذا المعامل لحساب الارتباط البسيط بين مجموعتين من ازواج الدرجات لمتغيرين x و y .

طريقة حساب معامل ارتباط بيرسون:

الطريقة الأولى:

وتتم باستخدام المعادلة التالية:

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \ \Sigma(y-\bar{y})^2}}$$

حيث ان :

$\sum(x-\bar{x})(y-\bar{y})$ مجموع حاصل ضرب الانحرافات.

$\bar{x}$ المتوسط الحسابي للمتغير x

$\bar{y}$ المتوسط الحسابي للمتغير y

$\sum(x-\bar{x})^2$ مجموع مربعات انحرافات قيم المتغير x عن متوسطها الحسابي.

$\sum(y-\bar{y})^2$ مجموع مربعات انحرافات قيم المتغير y عن متوسطها الحسابي.

n عدد ازواج القيم

مثال:

طبق اختبار الانبطاح المائل ثني الذراعين واختبار هارفارد للخطوة على عينة تتكون من 15 طالبا جامعيا، وقد حسبت درجات مؤشر اللياقة البدنية من اختبار هارفارد للخطوة، فكانت النتائج كالتالي:

الانبطاح المائل x	**28**	**23**	**13**	**16**	**17**	**24**	**25**	**26**	**10**	**33**	**25**	**20**	**18**	**16**	**36**
اختبار الخطوة y	**84**	**74**	**57**	**63**	**80**	**72**	**70**	**65**	**59**	**91**	**73**	**60**	**73**	**83**	**91**

المطلوب:

حساب معامل الارتباط بين نتائج الأداء في اختبار الانبطاح المائل ثني الذراعين ونتائج مؤشر اللياقة البدنية المحوب من نتائج اختبار هارفارد للخطوة، باستخدام معامل ارتباط بيرسون.

الحل:

الطلاب	**قيم x**	**قيم y**	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$x-\overline{x}$ $y-\overline{y}$
1	**28**	**84**	**6**	**11**	**36**	**121**	**66**
2	**23**	**74**	**1**	**1**	**1**	**1**	**1**
3	**13**	**57**	**−9**	**−16**	**81**	**144**	**144**
4	**16**	**63**	**−6**	**−10**	**36**	**60**	**60**
5	**17**	**80**	**−5**	**7**	**25**	**−35**	**−35**
6	**24**	**72**	**2**	**−1**	**4**	**−2**	**−2**
7	**25**	**70**	**3**	**−3**	**9**	**−9**	**−9**
8	**26**	**65**	**4**	**−8**	**16**	**−32**	**−32**
9	**10**	**59**	**−12**	**−14**	**144**	**168**	**168**
10	**33**	**91**	**11**	**18**	**121**	**198**	**198**
11	**25**	**73**	**3**	**0**	**9**	**0**	**0**
12	**20**	**60**	**−2**	**−13**	**4**	**26**	**26**
13	**18**	**73**	**−4**	**0**	**16**	**0**	**0**
14	**16**	**83**	**−6**	**10**	**36**	**−60**	**−60**
15	**36**	**91**	**14**	**18**	**196**	**252**	**252**

$N = 15$

$\sum x = 330$

$\sum y = 1095$

$\bar{x} = 22$

$\bar{y} = 73$

$\sum(x - \overline{x})^2 = 1714$

$\sum(y - \bar{y})^2 = 1714$

$\sum(x - \overline{x}) \quad / \quad \sum(y - \bar{y}) = 777$

يلاحظ ان درجات العمود x ودرجات العمودy تشمل قيم اختبار هارفارد للخطوة، وقد جلت النتائج المتعلقة بقيم الأزواج لكل طالب معا.

- نقوم بحساب المتوسط الحسابي لنتائج الطلاب في كل من الاختبارين فيكون متوسطهما كالتالي:

$$\bar{x} = \frac{330}{15} = 22$$

$$\bar{y} = \frac{1095}{15} = 73$$

- نقوم بحساب انحرافات قيم المتغير x عن متوسطها الحسابي وكذلك قيم المتغير y وتسجل الانحرافات في العمودين 4 و 5 مع تحديد الإشارة.
- نقوم بتربيع هذه الانحرافات برب كل قيمة في نفسها للتخلص من الإشارات السالبة، م يوضع الناتج في العمودين 6 و 7 .

- نقوم بحساب مجموع مربعات انحرافات المتغير x والمتغير y كل على حدة، مع وضع النتائج افل الجدول.
- نقوم بضرب انحرافات كل قيمة من قيم المتغير x، في انحرافات القيمة المناظرة لها للمتغير y، مع وضع الناتج في العمود الأخير رقم 8.
- نقوم بجمع حاصل رب الانحرافات قيم المتغير x في انحرافات قيم المتغير y مع وضع النتائج اسفل الجدول.
- تأتي النتائج وفق:

$$734 = \sum(x - \overline{x})^2$$

$$1714 = \sum(y - \bar{y})^2$$

$$777 = \sum(x - \bar{x})(y - \bar{y})$$

وبالتعويض في المعادلة:

$$\boldsymbol{r = \frac{777}{\sqrt{(734)(1714)}}}$$

$$\boldsymbol{= 0.69}$$

شروط استخدام معامل ارتباط بيرسون:

يعتبر معامل ارتباط بيرسون أكثر معاملات الارتباط دقة لكونه يستخدم جميع القيم المتجمعة من القياس، كما انه يدخل ضمن العديد من العمليات الإحصائية في مجال الإحصاء المعلمي، ويشترط عند استخدام هذا المعامل شرطان اساسيان هما:

- ان يتوزع المتغيران x, y في المجتمع الأصلي توزيعا معتدلا وهذا الشرط يقتضي بطبيعة الحال الا يكون التواء أحد المتغيرين ذا دلالة إحصائية.
- ان تكون العلاقة بين المتغيرين علاقة خطية، بمعنى انه اذا حسبت متوسطات درجات المتغير x ، ومتوسطات درجات المتغير y، فإننا نجد ان هذه المتوسطات تميل لان تقع على خطين مستقيمين، وان النقاط التي تعبر عن المتوسطات لا تنحرف كثيرا عن هذين المستقيمين، مما يدل على ان العلاقة بين المتغيرين علاقة خطية مستقيمة.

معامل التعيين:

من المعروف ان معال الارتباط المحسوب بين متغيرين لا يدل على نسبة، فنحن لا نستطيع ان نجزم ان معامل الارتباط الذي يساوي 0.80 بين المتغيرين يساوي مرتين مقدار الارتباط الذي يساوي 0.40 بين متغيرين اخرين مثلا.

من ناحية أخرى فانه لا يجوز ان نقرر ان معامل الارتباط الذي يساوي 0.80 يعكس 80% من مقدار العلاقة بين المتغيرين x,y مثلا وان معامل الارتباط الذي يساوي 0.40 يعكس 40% من مقدار العلاقة المحتملة بين المتغيرين الاخرين.

ويستخدم معامل التعيين للارتباط المحسوب بين أي متغيرين لتقدير نسبة التغاير بين هذين المتغيرين، حيث يشير مربع معامل الارتباط المحسوب بين هذين المتغيرين الى هذه النسبة.

فعند حساب معامل التعيين بالنسبة لمعاملي الارتباط 0.80، 0.40 فانه يلاحظ الاتي:

أن : r_{xy}= 0.80

وأن معامل التعيين = $(0.80)^2$

0.64=

وأن r_{ab}=0.40

وان معامل التعيين = $(0.40)^2$

0.16 =

حيث يمكن تفسير هذه النتائج على النحو التالي:

- ان معامل الارتباط الذي يساوي معامل تعيين يساوي 0.64 ، بمعنى انه للمتغيرين x,y 0.80 نسبة تباين مشترك تبلغ 0.64%.
- ان معامل الارتباط الذي يساوي 0.40، له معامل تعيين يساوي0.16، بمعنى انه للمتغيرين a,b نسبة تباين مترك تبلغ 16 %.
- انع كلما زادت القيمة لمعامل الارتباط المحسوب بين متغيرين، أدى ذلك الى زيادة معامل التعيين بين هذين المتغيرين.
- ان مضاعفة كعامل الارتباط بين أي متغيرين يؤدي الى زيادة التباين بينهما اربع مرات، فمعامل الارتباط الذي يساوي 0.80 له معامل تعيين يساوي 64%، ومعامل الارتباط الذي يساوي 0.40 له معامل تعيين يساوي 16% . حيث يلاحظ ان معامل التعيين للارتباط الأول يساوي 4 مرات معامل التعيين للارتباط الثاني.

معامل عدم التعيين:

وهو معامل يدل على مقدار التباين الخاص بكل متغيرين كل على حدة حيث يتم حابه باستخدام المعادلة التالية:

معامل عد التعيين = $1-(r)^2$

حيث أن:

R: معامل الارتباط المحوب بين المتغيرين

R^2: معامل التعيين.

فمعامل عدم التعيين في المثال السابق يمكن حسابه على النحو التالي:

$R_{xy} = 1-0.64 = 0.36$ أو 36%

$R_{ab} = 1-0.16 =$0.84 او 84%.

وهذه النتائج تعني:

- ان نسبة التباين المشترك بين المتغيرين x,y تساوي 64 % في حين تبلغ نبة التباين الخاصة بكل متغير على حدة 36%.
- ان نبة التباين المشترك بين المتغيرين a,b تساوي 16% في حين تبلغ نسبة التباين الخاصة بكل متغير على حدة 84 %.

معامل الاغتراب:

اشرنا الى ان معامل التعيين بين أي متغيرين يدل على نسبة مئوية للتباين المشترك بين هذين المتغيرين ، كما ان ه كلما زادت قيم معامل الارتباط بين المتغيرين زادت بالتالي قيم معامل التعيين التي تدل على قوة العلاقة بين المتغيرين0

ونظرا لعدم وجود معايير محددة تدل على ما اذا طان الارتباط المحسوب بين المتغيرين قويا ام عيفا، لذا يستخدم معامل الاغتراب كمؤشر للدلالة على الصفات او الخصائص التي لا يشترك فيها المتغيرين، فالاغتراب يقيس عكس ما يقيسه الارتباط، فهو مؤشر يدل على مدى اختفاء التلازم في التغير (التغير الارتباطي) في الظاهرة المقيسة.

معامل الاغتراب = $\sqrt{1 - r^2}$

حيث أن:

R^2 معامل التعيين بين المتغيرين

مثال:

إذا كان :

- معامل الارتباط بين الطول والوزن يساوي 0.91.
- معامل الارتباط بين الانبطاح المائل ثني الذراعين والشد للأعلى يساوي 0.96.
- ومعامل الارتباط بين التوازن والرشاقة يساوي 0.70.
- فانه يمكن حساب معاملات الاغتراب للارتباطات السابقة كالتالي:

معامل الاغتراب للارتباط الأول $= \sqrt{1 - 0.91^2}$

$0.41 =$

معامل الاغتراب للارتباط الثاني $= \sqrt{1 - 0.96^2}$

$= 0.28$

معامل الاغتراب للارتباط الثالث $= \sqrt{1 - 0.70^2}$

$= 0.71$

مما سبق نلاحظ الاتي:

- ان معامل الارتباط بين الطول والوزن الذي يساوي 0.91 تزيد قيمته عن معامل الاغتراب الذي يساوي 0.41 مما يؤكد ترابط العلاقة بين الطول والوزن كمتغيرين أكثر من اغترابهما.
- ان معامل الارتباط بين الانبطاح المائل ثني الذراعين والشد لأعلى الذي يساوي 0.96 تيد قيمته عن معامل الاغتراب الذي يساوي 0.28، مما يؤكد على ترابط العلاقة بين هذين المتغيرين كمقاييس لتحمل قوة منطقة الذراعين والمنكبين.
- واما بالنسبة لمعامل الارتباط بين التوازن والرشاقة والذي يساوي 0.70 فانه يلاحظ أنه يساوي تقريبا معامل الاغتراب الذي يساوي0.71.
- وقد اقر العلماء ان معامل الاغتراب الذي يساوي معامل الارتباط يمكن استخدامه كمحك لقبول معامل الارتباط المحسوب، وكحد أدني لوجود علاقة قوية بين المتغيرين موضوع البحث.

مؤشر التنبؤ:

يمدنا مؤشر التنبؤ بإمكانية التنبؤ الصحيح بالمتغير x عند معرفة ارتباطه بمتغير اخر y أي بمعنى اخر ما مقدار الدرجة المطلوب معرفتها بالنسبة للمتغير x حتى تساعدنا في معرفة المتغير y.

يستخدم لحساب مؤشر التنبؤ المعادلة التالية:

مؤشر التنبؤ = $\mathbf{1 - \sqrt{1 - r^2}}$

يستخدم كمقياس للتنبؤ بالمتغير التابع من المتغير المستقل، إذا كانت العلاقة الخطية بين المتغيرين معلومة، فاذا كان معامل الارتباط بين المتغيرين يساوي0.91، فانه يمكن تقدير مؤشر التنبؤ بالمتغير y من المتغير x كالتالي:

مؤشر التنبؤ $= 1 - \sqrt{1 - r^2}$

$= 1 - \sqrt{1 - 0.91}$

$= 1 - \sqrt{1 - 0.83}$

$= \sqrt{1 - 0.17}$

$= 0.59$

مؤشر التنبؤ يساوي 59%

وهذه النتيجة تعني ان معامل الارتباط بين المتغيرين والذي يساوي 0.91 يمكن استخدامه في التنبؤ باي من المتغيرين بدلالة المتغير الاخر بدرجة ثقة تساوي 59% .

الدلالة الاحصائية لمعامل الارتباط بيرسون:

تم الإشارة أعلاه الى ان حدود معامل الارتباط بين متغيرين تتراوح من (−1 الى +1) الا ان معامل الارتباط المحوب لا يكتسب دلالته من قيمته المطلقة، فالقيم المطلقة لمعامل الارتباط غير ذات أهمية مادام حجم العينة ودرجات الحرية المختلفة وقوانين الاحتمالات هي المحك الحقيقي لحساب دلالة معامل الارتباط.

تتأسس الدلالة الإحصائية لمعامل الارتباط على انه اذا كان معامل الارتباط المحسوب دال احصائيا، فان الدلالة في هذه الحالة تعني ان هذا المعامل يختلف عن قيمة الصفر، اما اذا معامل الارتباط المحسوب غير دال احصائيا فان هذا يعني ان معامل الارتباط لا يختلف عن قيمة الصفر.

من الملاحظات الجديرة بالاهتمام في هذا الخصوص ان المعلومات الممتدة من الدلالة الإحصائية لمعامل ارتباط بيرسون لا تساعد في تقدير درجة العلاقة التي تمتد بين متغيرات موضوع الدراسة، فعندما يزيد حجم العينة فانه يلاحظ ان معاملات الارتباط ذات القيم الصغيرة تصبح دالة احصائيا.

عادة ما تكون معاملات الارتباط مقبولة الدلالة اذا كانت عند مستوى 0.05 وهو مستوى يعني ان هذا المعامل يمكن ظهوره بين المتغيرين في 95 حالة من كل 100 حالة، كما تقبل بتقدير اكبر اذا كانت الدلالة عند مستوى 0.01.

معامل الارتباط سبيرمان:

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبيين، مثال ذلك قيا العلاقة بين تقديرات الطلبة في مادتين، او العلاقة بين درجة تفضيل المستهلك لسلعة معينة، ومستوى الدخل، فانه يمكن استخدام طريقة سبيرمان في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الاصلية، ويطلق على هذا المعامل معامل ارتباط سبيرمان. يتم استخدامه وفق المعادلة التالية:

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

حيث أن:

d: هي الفرق بين رتب مستويات المتغير الاولx، ورتب مستويات المتغير الثانيy، وعندما تكون هناك عدة مشاهدات بنفس الرتبة يعتبر الوسط الحسابي هو رتبة كل واحدة من تلك المشاهدات عند ترتيبها تصاعديا.

متطلبات تطبيق معادلة معامل ارتباط سبيرمان:

- أن توضع البيانات الاصلية في شكل رتب.
- ان يكون عدد ازواج الرتب 30 أو اقل.
- يتم استخدام الجداول الإحصائية، لتقويم الدلالة الإحصائية لمعامل الارتباط المحسوب، ويلاحظ ان هذه الجدول قد وضعت باعتبار عدد ازواج المشاهدات 30 فاقل.
- في حالة زيادة عدد ازواج المشاهدات عن 30 تستخدم المعادلة التالية لحساب قيمة معامل الارتباط بدلالة قيمة اختبار z التي تطابق مستوى 0.05 للطرفين.

$$r_s = \frac{z}{\sqrt{n-1}}$$

مميزات معامل الارتباط سبيرمان:

يعتبر من أكثر المعاملات شيوعا اذ انه:

- ويلة مناسبة لتحليل أنماط البيانات التي تكون في شكل رتب، حيث لا يصلح معامل الارتباط الخطي كوسيلة لتحليل مل هذه البيانات.
- يستخدم هذا المعامل لتقويم علاقات الارتباط غير الخطية. يتميز بسهولة وسرعة حساب نتائجه وذلك بالمقارنة بمعامل الارتباط الخطي بيرسون.

- يستخدم هذا المعامل حينما لا تتوفر المتطلبات الأساسية اللازمة لتطبيق الارتباط الخطي في الإحصاء البارومتري، فالارتباط الخطي يتطلب ضرورة ان تتوزع البيانات في المجتمعات الاصلية توزيعات معتدلة في حين لا يتطلب معامل ارتباط سبيرمان توافر هذا الشرط.

مثال:

فيما يلي تقديرات 10 طلاب في مادتي الإحصاء والمورفولوجيا

الاحصاء	أ*	ج*	د	د*	ب*	ج*	أ*	ب	ب*	ب*
المورفولوجيا	أ*	د	ج	ج	أ	ب	ب*	ب	ج	ب

المطلوب:

- حساب معامل الارتباط بين تقديرات الطلبة في المادتين.
- ما هو مدلوله؟

الحل:

- بفرض ان x هي تقديرات الإحصاء، وان y هي تقديرات المورفولوجيا فانه يمكن حساب معامل الارتباط بينهما باستخدام المعادلة:

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

الرتب	01	02	03	04	05	06	07	08	09	10
تقديرات الاحصاء	أ*	أ	ب*	ب*	ب*	ب	ج*	ج*	د*	د
رتب x	1	2	3+4+5/3=4			6	7+8/2=7.5		9	10
تقديرات المورفولوجيا	أ*	أ	ب*	ب	ب	ب	ج	ج	ج	د
رتبy	1	2	3	4+5+6/3=5			7+8+9/3=8			10

إذن يمكن حساب المجموع $\sum_d 2$ كما يلي:

x	Y	رتب x	رتب y	D	d^2
أ	أ*	2	1	1	1
ج*	د	7.5	10	−2.5	6.25
د	ج	10	8	2	4
د*	ج	9	8	1	1
ب*	ا	4	2	2	1
ج*	ب	7.5	5	2.5	6.25
ا*	ب*	1	3	−2	4
ب	ب	6	5	1	1
ب*	ج	4	8	−4	16
ب*	ب	4	5	−1	1
					44.5

$\sum_d 2$ = 44.5

معامل الارتباط هو:

نقوم بحساب معامل ارتباط سبيرمان بتطبيق المعادلة :

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

$$r_s = 1 - \frac{6(44.5)}{10(10^2 - 1)}$$

$$= 1 - \frac{267}{990}$$

$$= 1 - 0.2697$$

$$= 0.7303.$$

مدلول معامل الارتباط:

بما ان $r = 0.703$ فان ذلك يدل على وجود ارتباط طردي قوي بين تقديرات الطالب في مادة الإحصاء ومادة المورفولوجيا.

مثال:

في تقييم لنتائج عينة تتكون من 11 طالبا في امتحاني الكرة الطائرة وكرة السلة، كانت نتائج التقييم كما في الجدول التالي:

تسلل الطالب	تقييم نتائج كرة الطائرة	تقييم نتائج كرة السلة
1	جيد	ضعيف جدا
2	ضعيف	ممتاز
3	مقبول	ممتاز
4	جيد	جيد
5	ممتاز	مقبول

6	مقبول	جيد جدا
7	ضعيف جدا	مقبول
8	جيد جدا	ضعيف
9	ممتاز	جيد
10	ضعيف	ضعيف جدا
11	جيد جدا	مقبول

المطلوب: إيجاد العلاقة بين أداء الطالب في الامتحانين باستخدام معمل ارتباط سبيرمان

الحل:

- يتم إعطاء رتبة كل طالب حب مستوى النتائج او التقييم في كلا الامتحانين امتحان الكرة الطائرة وامتحان كرة السلة.

التسلسل	X1	X2	di	di^2
1	6.5	1.5	5	25
2	2.5	10.5	−8	64
3	4.5	10.5	−6	36
4	6.5	7.5	−1	1
5	10.5	5	5.5	30.5
6	4.5	9	−4.5	20.25
7	1	5	−4	16
8	8.5	3	5.5	30.25
9	10.5	7.5	3	9
10	2.5	1.5	1	1
11	8	5	3.5	12.25

				$\sum d^2$=245	$\sum d$=0.0

وبتطبيق صيغة معامل الارتباط التالية نحصل على:

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

$$r_s = 1 - \frac{6\ (245)}{11(120)}$$

$$r_s = -0.114$$

أي ان العلاقة بين نتيجتي الامتحان سالبة وضعيفة.

مميزات استخدام معامل الارتباط سبيرمان:

لاستخدام معامل الارتباط سبيرمان نتبع الخطوات التالية:

- اذا كانت البيانات موزعة في هيئة رتب، نقوم بحساب عدد أزواج الرتب، وفي الحا لات التي تكون فيها البيانات غير موضوعة في شكل رتب نقوم بتحويل بيانات المتغير الأول الى رتب من 1 الى n وكذا المتغير الاني، ثم نقوم بعد ذلك بحصر عدد ازواج الرتب لنحصل على n .
- حساب الفروق بين كل زوج متناظر من الرتب، ويتم ذلك بطرح الرتبة الأقل من الرتبة الأكبر.
- نقوم بتربيع كل فرق من هذه الفروق كل على حدة، يلي ذلك إيجاد حاصل جميع هذه المربعات فنحصل على $\sum d^2$.
- نقوم بحساب معامل ارتباط سبيرمان بتطبيق المعادلة، فنحصل من خلالها على إحصاء العينة.

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

- اذا كان حجم العينة يساوي او اصغر من 30 نقوم بإيجاد القيمة الحرجة الجدولية ل r_s.

مثال:

اوجد معامل ارتباط الرتب للبيانات التالية:

x	25	20	30	45	10	17	23	21
y	70	47	73	40	45	50	48	63

الحل: يكون ترتيب القيم y , x على النحو التالي:

ترتيب x	3	6	2	1	8	7	4	5
ترتيب y	2	6	1	8	7	4	5	3

تتم بقية الحسابات من خلال الجدول التالي:

ترتيب x	3	6	2	1	8	7	4	5	مج
ترتيبy	2	6	1	8	7	4	5	3	
di	1	0	1	7	1	3	1	2	
Di^2	1	0	1	49	1	9	1	4	66

بتطبيق المعادلة أعلاه نجد:

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

$$r_s = 1 - \frac{6\ (66)}{8(8^2 - 1)}$$

$$r_s = 0.21$$

بما ان $r = 0.21$ فيمكننا القول ان الارتباط ضعيف.

وكما ذكرنا سالفا فان هذه الطريقة تستخدم في حالة ما إذا كانت البيانات غير متساوية، اما إذا كان البعض منها متساو فإننا نتبع الخطوات التالية:

- نرتب القيم ونعطي لكل قيمة ترتيبا، كما لو ان لي فيها قيما متساوية بحيث أكبر قيمة تأخذ الرتبة 1 والقيمة الأقل تأخذ الرتبة الأخيرة.
- نوجد الوسط الحسابي لرتب كل مجموعة من البيانات المتساوية، ونعطي لكل قيمة من القيم المتساوية ترتيبا يساوي هذا الوسط.

مثال: اوجد معامل ارتباط الرتب للبيانات التالية:

ترتيب x	**25**	**20**	**30**	**45**	**20**	**17**	**20**	**25**
ترتيب y	**70**	**45**	**73**	**40**	**45**	**50**	**45**	**63**

نلاحظ انه توجد قيم متساوية في كل من المتغيرين، لذلك نقوم بإيجاد متوسط الرتب للقيم المتساوية بعد ترتيبها، ونعطي كل قيمة من القيم المتساوية رتبة تساوي هذا المتوسط وذلك عند موقعها الأصلي كما هو موضح في الجدول التالي.

بالنبة للمتغير x، مثلا نجد ان القيمة 20 مكررة 3 مرات ترتيبها على التوالي 5،6،7 لذلك يكون متوسط هذه الرتب هو (5+6+7)/3= 6.

لذلك نعطي قيمة تساوي 20 عند موقعها في الجدول الأصلي ترتيبا يساوي 6، ونقوم بنفس الطريقة بالنسبة للقيم المتساوية الأخرى.

ترتيب x	3.5	6	2	1	6	8	6	3.5	مج
ترتيبy	2	6	1	8	6	4	6	3	
di	1.5	0	1	7−	0	4	0	0.5	
Di^2	2.25	0	1	49	0	16	0	0.25	68.5

وبتطبيق المعادلة:

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

$$r_s = 1 - \frac{6\ (68.5)}{8(8^2 - 1)}$$

$$= \quad \mathbf{0.18}$$

معامل الارتباط كندال للرتب:

يهدف معامل الارتباط كندال الى قياس العلاقة بين متغيرين كلاهما رتبي ويرمز له بالرمز Taw ، يستخدم في مجال الاحصاء اللاباروميتري كمقياس للاتفاق بين مجموعتين منفصلتين من الرتب وفقا لمجموع التغيرات في الوضع.

يتطلب معامل الارتباط كندال نفس الشروط التي يتطلبها معامل ارتباط سبيرمان ، فهو تطبيق لحساب معامل الارتباط بين انين يتم التعبير عنها بمقاييس الترتيب، حي يمدنا معامل كندال كميا بمقدار العلاقة بين هذين المتغيرين.

يكمن الفرق الأساسي بين معامل كندال ومعامل سبيرمان في ان الأول يتطلب حاب انعكاسات (مقلوبات) الرتب بدلا من معالجة الرتب ذاتها كبيانات عند حساب الارتباط كما هو متبع في معامل سبيرمان، فمعامل كندال يعتمد فقط على عدد المقلوبات لكل زوج من المشاهدات او الافراد في فئتين من فئات الترتيب.

المعادلة:

$$\text{ر ك} = \frac{(\text{ع}-\text{م})}{\mathbf{1/2}\ \text{ن}\left(\text{ن}-\mathbf{1}\right)}$$

حيث:

ر ك : هو معامل كندال.

ع= العدد الكلي للرتب الكبيرة.

م= العدد الكلي للرتب الصغيرة.

ن: عدد المشاهدات او الدرجات التي تم ترتيبها.

مثال:

لنفترض ان أحد الباحثين قام بتطبيق اختبارين أ، ب على 12 طالبا، وقد حصل على النتائج الخام المبينة بالجدول التالي:

الطلاب	أ	ب	ج	د	هـ	و	ز	ح	ط	ي	ك	ل
الاختبار أ	20	26	17	16	15	23	22	24	19	28	30	10
الاختبار ب	70	80	40	45	38	49	77	76	72	47	36	53

المطلوب: حساب معامل كندال لرتب هاتين المجموعتين؟

الحل:

نقوم بوضع رتب كل مجموعة من النتائج أ و ب كل على حدة بحيث تعطى اكبر درجة للرتبة1، والدرجة الثانية لها الرتبة، وهكذا وفقا لما هو مبين في الجدول التالي:

الطلاب	أ	ب	ج	د	هـ	و	ز	ح	ط	ي	ك	ل
الاختبار أ	7	3	9	10	11	5	6	4	8	2	1	12
الاختبار ب	5	1	9	8	10	6	2	3	4	74	11	12

نقوم بإعادة ترتيب رتب اختبار أ بحيث تظهر في كل ترتيب طبيعي متسلسل من 1، 2، 3....12، مع ملاحظة إعادة ترتيب الرتب الخاصة بالاختبار ب وفقا للترتيب الذي تم بالنسبة للاختبار أ.

الطلاب	أ	ب	ج	د	هـ	و	ز	ح	ط	ي	ك	ل
الاختبار أ	1	2	3	4	5	6	7	8	9	10	11	12
الاختبار ب	11	7	1	3	6	2	5	4	9	8	10	12

نبدأ بالرتب على الاختبار ب من الجهة اليمنى للجدول، فنأخذ الرتبة 11ونحسب عدد الرتب التي تقع فوقها والرتب التي تقع تحتها، فنجد ان الرتب التي تقع فوقها (اكبر منها) تساوي 1 وهي الرتبة 12، ومجد ان الرتب التي تقع تحتها تساوي 10، وهي الرتب 1،2،3،4،5،6،7،8،9،10 ويلاحظ ان جميع هذه الرتب تقع على يسار الرتبة 11 بالنسبة للاختبار ب.

نبدأ بعد ذلك بالرتبة التالية لنفس الاختبار وهي الرتبة 7، وبنفس الطريقة السابقة نعد الرتب الأكبر منها فنجد انها تساوي4، والرتب الأصغر منها فنجد انها تساوي 6.

وبالمثل نأخذ الرتبة التالية 1 للاختبار ب فنجد ان الرتب الأكبر منها وتقع على يسارها تساوي9، وان الرتب الأقل منها وتقع على يسارها تساوي صفرا حيث لا توجد أي رتبة اقل من1.

وبالنسبة للرتبة 10 فنجد ان عدد الرتب الأكبر منها التي تقع على يسارها يساوي 1، وان عدد الرتب الأقل منها والتي تقع على يسارها يساوي صفرا، وبالنسبة للرتبة الأخيرة، فنجد ان عدد الرتب الأكبر والاقل منها والتي تقع على يسارها يساوي صفرا.

بعد الانتهاء من الاجراء الموضح سلفا بالنسبة لكل الرتب في الاختبار ب، نقوم بوضع النتائج التي يتم التوصل اليها في الجدول الاتي:

عدد الرتب التي تزيد عليها ع	**1**	**4**	**9**	**7**	**4**	**6**	**4**	**4**	**2**	**2**	**1**	**0**
رتبة المتغير ب	**11**	**7**	**1**	**3**	**6**	**2**	**5**	**4**	**9**	**8**	**10**	**12**
عدد الرتب التي تقل عنها م	**10**	**6**	**0**	**1**	**3**	**0**	**1**	**0**	**1**	**7**	**0**	**0**

نقوم بحساب قيمة كل من ع، م كالتالي:

ع= 1+4+9+7+4+6+4+4+2+2+1 =44

م= 1+6+1+3+1+1+1=22

(ع- م)= 44-22=22

وهي القيمة التي نستخدمها للكشف عن الدلالة الإحصائية للارتباط من خلال الجداول الإحصائية.

وبما ان :

(ع−م)= 44−22=22

ن= 12

اذن نقوم بالتعويض في المعادلة:

رك = 22/ 2/1*12(12−1)

= 0.333

الدلالة الإحصائية لمعامل ارتباط كندال:

تقوم الدلالة الإحصائية لمعامل ارتباط كندال على أساس ان المتغيرات التي تقف ورا الرتب تتوزع توزيعا متصلا، ويتأسس اختبار هذه الدلالة على الفرض الصفري الذي يقرر ان هذا المعامل يساوي صفرا.

المراجع:

- الشربيني، زكريا أحمد، دمنهوري، رشاد بن صالح، مطحنة، السيد خالد .(2019) . الإحصاء الوصفي والاستدلالي في علم النفس والتربية والاجتماع (ط.1). دار روابط للنشر وتقنية المعلومات.
- العدل، عادل محمد .(2022) .الإحصاء النفسي والتربوي: تطبيقات على SPSS (ط.1). دار الكتاب الحديث.
- البلداوي، عبد الحميد عبد المجيد .(2007) .أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع وتحليل البيانات يدويًا وباستخدام برنامج SPSS (ط.1). دار الشروق للنشر والتوزيع.
- حراوبية، ليندة .(2017) .مدخل إلى الإحصاء الوصفي .ديوان المطبوعات الجامعية.
- راتول، محمد .(2018) .الإحصاء الوصفي .ديوان المطبوعات الجامعية.
- رضوان، محمد نصر الدين .(2002) .الإحصاء الوصفي في علوم التربية البدنية والرياضية (ط.1). دار الفكر العربي.