تقديم

خطوات منهجية أساسية

مفهوم العية الاحصائية وأنواعها وطرق إحصائها

أنواع المتغيرات

مقاييس النزعة المركزية:

المتوسط الحسابي

المنوال

الوسيط

مقاييس التشتت:

المدى

التباين

الانحراف المعياري

معامل الاختلاف

أنواع العلاقة بين المتغيرات

معاملات الارتباط:

معامل ارتباط برسون Coefficient de corrélation de Pearson

معامل الارتباط سيرمن Coefficient de corrélation de spearman

معامل الارتباط الثنائي.

اختبارات الفروق:

اختبار أروين فخنر لعينتين مستقلتين .

اختبار Z

اختبار T

اختبارالتباين

اختبار حسن المطابقة كا2.

المعالجة الاحصائية للبيانات اعتمادا على برنامج SPSS:

إدخال البينات وترميزها:

إدخال بينات المتغيرات النوعية

إدخال بينات المتغيرات الكمية

تطبيق نماذج من المقاييس الإحصائية باستخدام برنامج SPSS:

مقاييس النزعة المركزية

مقاييس التشتت

•

تقديم:

بدأ علم الإحصاء باكتساب أهمية متعاظمة خاصة مع ظهور وتطور مفهوم وأعمال الدولة الحديثة، فقد كان يعني بجمع وتنظيم القضايا المتعلقة بشؤون الدولة، إلا أنه أصبح يقدم طرقاً مترابطة مبنية على أساس منطقي. بجمع البيانات

حول الظواهر وتلخيصها وتحليلها وصولاً إلى تقديم استنتاجات حول البيانات التي تم جمعها وتوسيع نطاق تلك الاستنتاجات إلى مدى أوسع من حدود تلك البيانات وتجدر الإشارة إلى أن بلورة علم الإحصاء بملامحه الحالية بدأت منذ نهاية القرن الثامن عشر واشتغل في تطويره آنذاك بعض العلماء مثل لابلاس (Laplace) وجاوس (Gauss).

ونظراً للاهتمام المتزايد بهذا العلم فقد توسعت طرقه وأساليبه واعتمدت على نماذج رياضية ومعادلات تحدد العلاقة بين المتغيرات وأصبح يضم عدة فروع منها الإحصاء التطبيقي والإحصاء الرياضي والاحتمالات.

هذا وقد استفادت العديد من العلوم، مثل علم الاجتماع وعلم النفس والجغرافيا وغيرها، كما تطورت أخرى بسبب استخدامها أساليب علم الإحصاء، فبعد أن كانت هذه العلوم تعتمد على الأسلوب الوصفي والإنشائي أخذت تعتمد على الأسلوب الكمي الذي ساهم في فهم طبيعة العلاقات التي تحكم الظواهر، وتعدى ذلك إلى تقديمه وسائل تساعد على التنبؤ بسلوك الظواهر والتحكم بها والسيطرة عليها في ضوء ظروف معينة.

خطوات الطريقة الإحصائية:

تتطلب الدراسة الإحصائية عادة إجراءات تبدأ بجمع البيانات والمعلومات عن الظاهرة موضوع الدراسة يليها عملية تبويب هذه البيانات وعرضها على شكل جداول أو رسومات بيانية ثم تحليل هذه البيانات بالطرق المناسبة من أجل استخلاص النتائج والتعميمات واتخاذ القرارات أو وضع التوصيات والاقتراحات.

ويمكن تلخيص أهم خطوات الطريقة الإحصائية كما يلي:

1- تحديد هدف الدراسة الإحصائية أو مشكلة البحث بشكل دقيق كي نتمكن من تحديد ماهية البيانات المطلوبة.

2- صياغة الفرضيات باعتبارها عبارات محددة حول معلم من معالم المجتمع حيث نرغب بفحص الفرضية بهدف قبولها أو رفضها.

3- تحديد مجتمع الدراسة ووحدة المجتمع أي تحديد ذلك الإطار الذي يحتوي على كافة المفردات التي تشملها الدراسة الإحصائية، فمثلاً يكون مجتمع الدراسة لعملية التعداد السكاني للملكة هم جميع سكان المملكة ووحدة المجتمع هي أي فرد من السكان. أما إذا أردنا تحديد متوسط عمر مصابيح مصنع ما فإن المجتمع هو المصابيح التي ينتجها المصنع ووحدة المجتمع هو المصباح. وكذلك يكون متوسط مجتمع الدراسة التي تهدف إلى تحديد متوسط أعمار طلبة المرحلة الأساسية في مدارس محافظة العاصمة هم جميع الطلاب والطالبات المنتظمين في مدارس المحافظة وقت إجراء الدراسة، ووحدة المجتمع هي الطالب. وينبغي أن يتم إجراء تصميم للتجربة للمساعدة في تحديد طريقة جمع البيانات.

4- جمع البيانات الإحصائية ويتم ذلك بتحديد مصادر البيانات والمعلومات اللازمة للظاهرة أو الموضوع المراد دراسته ومن ثم تحديد أسلوب جمع هذه البيانات في ضوء طبيعة البيانات المطلوبة وحجم ونوع مجتمع الدراسة، والإمكانات المادية والزمنية والفنية المتوفرة للباحث. وتقسم مصادر جمع البيانات إلى:

أ- **البيانات المباشرة**: التي يجمعها الباحث عن مفردات الظاهرة والتي يحصل عليها عادة عن طريق الميدان مباشرة وذلك عن طريق القيام بها بنفسه بالعد أو القياس للوحدات المطلوبة، أو عن طريق الأفراد الذين يتم إعدادهم وتدريبهم لهذه الغاية ويتم بالمقابلات الشخصية وتعبئة الاستمارات الإحصائية التي يعدها لهذه الغاية ويصممها لغرض الدراسة وهي تحتوي على مجموعة من الأسئلة تهدف للاستفسار عن البيانات المطلوبة وتعد حسب الأصول المعروفة لأعداد الاستمارات.

ب- **البيانات غير المباشرة**: والتي يحصل عليها الباحث عن طريق السجلات أو الوثائق التاريخية أو الجداول الإحصائية المنشورة والتي تتعلق بموضوع الدراسة وتكون تحت تصرف الباحث دون الحاجة إلى جمع البيانات من مفردات المجتمع الإحصائي مباشرة. وكمثال على ذلك النشرات الإحصائية التي تصدرها دائرة الإحصاءات العامة من خلال التعداد العام للسكان والمساكن وإحصاءات التجارة الخارجية وغيرها.

وتقسم أساليب جمع البيانات إلى:

أ- أسلوب الحصر الشامل: أي الحصول على البيانات من كافة أفراد المجتمع الإحصائي.

ب- أسلوب العينة: أي الحصول على البيانات من مجموعة جزئية من أفراد مجتمع الدراسة تسمى بالعينة، تختار بأحد الطرق الإحصائية المناسبة. شريطة أن تتوفر فيها خصائص المجتمع الأصلي أي أن تكون ممثلة لأفراد مجتمع الدراسة تمثيلاً تاماً، بحيث تمكن الباحث من تعميم النتائج المتحصلة من بيانات العينة على المجتمع الإحصائي بالكامل. وللعينة الجيدة شروط ومواصفات وطرق اختيار

تختلف باختلاف غرض الدراسة وطبيعة مجتمع الدراسة وإمكانات الباحث. إن أسلوب العينة هذا هو الأسلوب الأكثر شيوعاً من أسلوب الحصر الشامل لإنه يوفر من الوقت والجهد ومن التكاليف المادية، ويعطي نتائج جيدة في معظم البحوث إذا روعيت الشروط وإجراءات التصميم لهذا الأسلوب.

5- فرز وتنظيم وتبويب وتصنيف البيانات الإحصائية.

6- عرض البيانات الإحصائية الواحدة أو أكثر من طرق العرض المناسبة من الرسومات البيانية أو الأشكال الهندسية أو غيرها.

7- دراسة وتحليل البيانات الإحصائية باستخدام الطرق الإحصائية كمقاييس النزعة المركزية ومقاييس التشتت والارتباط أو السلاسل الزمنية والأرقام الهندسية وغيرها.

8- اختبار الفرضيات واستخلاص النتائج نتيجة الدراسة والتحليل واتخاذ القرارات اللازمة التي من أجلها أجريت الدراسة.

أنواع الاحصاء:

يرى بعض الإحصائيين، مثل: جون تاكي (John Tuky) بأن هناك نوعين من الإحصاء : الإحصاء الاستثماري(Statistique exploratoire) والإحصاء التأكيدي(Statistique confirmatoire). ويعني هذا أننا نستثمر المعطيات الإحصائية ومبياناتها في عملية المعالجة. ثم، نلتجئ إلى الإحصاء التأكيدي لتثبيت فرضية ما أو تفنيدها.

ويشبه هذا التقسيم ما تعارف عليه الدارسون حينما يشيرون إلى إحصاء وصفي وإحصاء استنتاجي أو استدلالي. فالإحصاء الوصفي هو تجميع البيانات، وتلخيصها، ومعالجتها، وتحويلها إلى معلومات من أجل الوصول إلى الخصائص الأساسية لمجتمع إحصائي. ونستخدم هذا النوع من الإحصاء عندما يكون حجم مجتمع الدراسة صغيرا أو بسيطا، حيث يمكننا تغطية جميع آراء ذلك المجتمع (دراسة قسم عدد تلاميذه 25 تلميذا). ويعرف الإحصاء الوصفي أيضا بأنه «مجموعة من الأساليب الرياضية التي تجعل بيانات موضوع البحث ومعلوماته تبدو أكثر وضوحا ودقة ودلالة. ولا يؤدي هذا النوع من الإحصاء إلى إصدار الأحكام أو إلى التنبؤ. فهو يساعد فقط على إجراء عمليات التنظيم والتبويب والوصف والتحليل.»[6]

ويتضمن الإحصاء الوصفي مجموعة من الأساليب والآليات الإجرائية، مثل: تنظيم البيانات، ومقاييس النزعة المركزية، ومقاييس التشتت، ومقاييس العلاقة، ومقاييس الوضع النسبي.

أما الإحصاء الاستنتاجي ، فهو ذلك الإحصاء الذي يهتم بتحليل البيانات، واستخدام النتائج، ثم تفسيرها، واستعمالها لاتخاذ القرارات الممكنة، ويسمى أيضا بعلم القرارات. ويبدأ حين ينتهي الإحصاء الوصفي. ويعني هذا أن الإحصاء الاستنتاجي هو «عبارة عن مجموعة من الأساليب الرياضية التي تسمح للباحث بالاستنتاج والتفسير والتحليل، وبإصدار الأحكام العلمية على نتائج البحث، ثم القدرة على التنبؤ بالظواهر والأحداث.[7]»

هذا، ويتميز الإحصاء الاستنتاجي أو الاستدلالي عن الإحصاء الوصفي بخاصيتين مميزتين هما: التنبؤ وإصدار الحكم[8]. ويعتمد على نظرية الاحتمالات في استقراء النتائج، و اتخاذ القرارات المناسبة بخصوص المجتمع من خلال العينة.

هذا، ويتضمن الإحصاء الاستنتاجي الأساليب التالية، مثل: مقاييس الدلالة، والدلالة الإحصائية للفوارق بين الإحصائيات، والدلالة الإحصائية لمعاملات الارتباط، ومقاييس مربع « كاي»، وتحليل التباين.

ومن باب العلم، فقد بدأ الإحصاء تاريخيا وصفيا، لينتقل – فيما بعد– إلى إحصاء استنتاجي استدلالي لاتخاذ القرارات الصائبة.

أنواع المتغيرات:

يتضمن مصطلح متغير شيئاً يتغير، ويأخذ قيماً مختلفة أو صفات متعددة؛ فهو مفهوم يعبر عن الاختلافات بين عناصر فئة معينة مثل: الجنس (ذكر، أنثى)، التحصيل (جيد، متوسط، ضعيف) الحالة الاجتماعية (متزوج، أعزب، مطلق) الانتماء السكني (ريفي، حضري).

وإذا كنا نعلم أنه لا توجد صعوبة في قياس وزن الأطفال أو عدد الناجحين في امتحان ما أو عدد أفراد في مؤسسة من المؤسسات، فإن هناك من المتغيرات ما يصعب علينا قياسه كمتغير تقدير الذات أو الشعور بالضغط المهني أو الاتجاه نحو عمل المرأة أو الذكاء أو المهارات والحوافز والتعلم، فكلها متغيرات تنطوي على صعوبة في قياسها.

المتغير إذاً مصطلح يدل على صفة محددة، تأخذ عدداً من الحالات أو القيم ويشترط فيه قابليته للقياس، و يعبر المتغير أيضا عن مفهوم معين يجرى تعريفه إجرائياً في ضوء إجراءات البحث، ويتم قياسه كمياً أو وصفه كيفياً، فالذكاء مثلاً صفة عقلية يمتلكها الأفراد بدرجات متفاوتة، وهو بذلك متغير لأنسه لا يوجد بنفس القيمة أو الدرجة أو المستوى عند جميع الأفراد.

إن الأصل في المتغيرات هو الاختلاف والتنوع، وهو ما يقصد بـه تنـوع مستويات المتغير سواء كان ذلك حسب قيمها الرقمية مثل: الوزن، التحصيل الدراسي...الخ، أو حسب مستوياتها الاسمية مثل نوع البيئة: حضر، ريف، أو نوع الجنس: ذكر، أنثى أو نوع الوظيفة: مدرس، إداري...الخ). وقد تكون للمتغير عدة مستويات، فعندما نقول مثلاً الجنس، فلهذا المتغير مستويين هما: ذكر أو أنثى وكل واحد منهما يمثل خاصية واحدة فقط، وعندما نتحدث عن متغير التعليم فلا نستطيع تجريده من مستوياته المختلفة المتمثلة في المرحلة الابتدائية والمرحلة المتوسطة

والمرحلة الثانوية.... فالتعليم هو المتغير في حين نقصد بالمستويات مجال الاختلاف لذلك المتغير، ومن هنا لا يمكن أن تمثل أي مفردة خاصيتين في آن واحد، أي لا يمكن أن نصف مفردة بأكثر من وصف في متغير واحد، فالمتغير هو أي مستوى يمثل خاصية وصفية واحدة.

أنواع المتغيرات:

فيما يلي نقدم جدولا نلخص فيه أهم أنواع المتغيرات وخصائص كل نوع:

نوع المتغير		الخصائص
كمي	متصل	- يقاس باستخدام وسائل القياس من مستوى المسافة، حيث تمثل قيم المتغيرات فروقاً في الدرجة على متصل واحد هو متصل المتغير - يتكون من الأعداد الصحيحة والكسور ومن أمثلته الذكاء، التحصيل. - لا توجد فيه فجوات بين قيمه.
	منفصل	- جميع قيمه صحيحة مثل: عدد أفراد الأسرة.
نوعي (اسمي)		- وظيفته هي تصنيف المفهوم في فئات، مثل النوع (ذكر، أنثى) (ناجح، راسب). - الأرقام في هذه المتغير لا تعبر عن كميات من خصائص فالاختلاف هنا ليس في الدرجة وإنما في النوع.
مستقل		- يستعمل في البحوث التجريبية أو شبه التجريبية، إنه المتغير التجريبي الذي يعالجه الباحث ليرى أثره على المتغير التابع وهو غالبا ما يكون تصنيفياً.
تابع		- يَظهر عليه أثر المتغير المستقل. - غالباً ما يكون متصلا.

مفهوم العينة الإحصائية وأنواع العينات وطرق اختيارها:

العينة الإحصائية هي جزء من المجتمع الإحصائي المدروس أي مجموعة جزئية من أفراد المجتمع الإحصائي يتم اختيارها لأغراض جمع البيانات وإصدار الأحكام واتخاذ القرارات. والمجتمع الإحصائي يختلف باختلاف نوع وطبيعة الدراسة الإحصائية المطلوبة حيث يعني، كما أسلفنا، المجال الذي يهتم الباحث بدراسته والتعرف على خواصه ومحاولة التنبؤ بصفاته ومميزاته الهامة، أي هو الإطار الذي يحتوي جميع المفردات التي يشملها موضوع الدراسة الإحصائية.

وبالرغم من أن الدراسة التي تأخذ البيانات المطلوبة من كافة مفردات المجتمع الإحصائي تعطي نتائج أكثر تمثيلاً، وهي عملية واسعة ومكلفة مادياً وتتطلب وقتاً وجهداً، خاصة عندما يكون المجتمع كبيراً أو غير متوفر بالكامل، أو إن دراسته ستصيب عناصره بالفساد والخسارة، فإن الأسلوب الأكثر عملية وتلجأ إليه معظم الدراسات الإحصائية، هو أسلوب العينة، أي أخذ البيانات من جزء ممثل تمثيلاً تاماً

لأفراد مجتمع الدراسة، والنتائج التي نحصل عليها من بيانات العينة، نعممها على كافة مفردات مجتمع الدراسة.

ومن أهم أنواع العينات الاحتمالية ما يلي:

1 . العينة العشوائية البسيطة (Simple Random Sample):

وهي تلك العينة التي تسحب من مجتمع الدراسة بحيث يكون احتمال (فرص) ظهور أية مفردة من مفردات المجتمع الإحصائي في العينة متساوياً، وبمعنى آخر تعني إعطاء كل فرد من المجتمع نفس الفرصة للظهور في العينة. ويتم اختيارها كما يلي:

أ. طريقة البطاقات أو القرعة: أي سحب عدد من البطاقات يناظر حجم العينة من صندوق يحتوي على جميع مفردات المجتمع.

وعليه يكون عدد العينات (n) من مجتمع حجمه (N):

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

مثال:

إذا كان حجم عينة n = 2 من أصل حجم مجتمع N = 3 فإن عدد العينات الممكنة:

$$\frac{!3}{!(2-3)!2} = \binom{3}{2}$$

$$\frac{1\times 2\times 3}{1\times 1\times 2} =$$

= 3 عينات ممكنة

ويشير الرمز (!) إلى مضروب العدد، فعلى سبيل المثال مضروب العدد 4 ويرمز له بالرمز 4 ! يساوي 4 × 3 × 2 × 1 = 24.

مثال:

إذا كان لدينا (5) طلاب وأردنا اختيار طالبين عشوائياً، فما هو عدد الطرق الممكنة لاختيار الطالبين؟

إن عدد الطرق الممكنة هو: $\binom{5}{2} = \frac{5!}{2!\times(5-2)!}$

$= \frac{1\times2\times3\times4\times5}{1\times2\times\cancel{3}\times1\times2} = 10$ طرق مختلفة.

بمعنى أنه يوجد عشر بطاقات يكتب عليها أسماء طالبين ويتم اختيار بطاقة من العشر بطاقات عشوائياً.

فإذا كانت أسماء الطلاب a, b, c, d, e فإن العشر بطاقات يكون مكتوب عليها: ab, ac, ad, ae, bc, bd, be, cd, ce, de ويتم سحب أي بطاقة من العشر بطاقات عشوائياً.

ب. استخدام جداول الأرقام العشوائية:

مثال:

لسحب عينة مكونة من (5) أفراد في مجتمع يحتوي على (500) شخص (باستخدام جدول الأرقام العشوائية) تتبع الخطوات التالية:

نرقم الأفراد أرقاماً متسلسلة من (001)، (002)، (003)، إلى (500) ومن ثم نختار (5) أرقام من جدول الأرقام العشوائية ونبدأ العد عامودياً أو أفقياً ويعتبر أي رقم يقل عن أو يساوي 500 مقبول كعنصر في العينة إلى أن نصل للعدد المطلوب وهو 5 أفراد وبالنظر إلى جدول الأرقام العشوائية في الملحق يكون الأفراد هم ذوي الأرقام 317، 309، 228، 418، 436 عندما نبدأ من اليسار.

2- العينة المنتظمة (Systematic Sampling)

يتم اختيارها من خلال ترقيم عناصر المجتمع الإحصائي بحيث يتم تحديد قاعدة للاختيار تستند على تحديد اختيار العنصر الأول فإذا أردنا اختيار 5 طلاب من مجتمع

من الطلاب يتألف من 50 طالب فإن قاعدة الاختيار تكون على أساس حساب حاصل قسمة $\frac{\text{حجم المجتمع}}{\text{حجم العينة}}$ وفي مثالنا فإن $\frac{\text{حجم المجتمع}}{\text{حجم العينة}}$ =10 وعليه فإننا سنختار طالب واحد من بين كل عشرة طلاب وبذلك نقوم باختيار الطالب الأول من خلال اختيار عدد يتراوح بين (0–9) عشوائياً فإذا حصلنا على الرقم 4 بطريقة عشوائية فإن الطالب رقم 4 يكون الطالب الأول في العينة ثم نقوم بزيادة حاصل قسمة $\frac{\text{حجم المجتمع}}{\text{حجم العينة}}$ الذي يساوي 10 إلى الرقم 4 فنحصل على 14 فيكون هو الطالب الثاني ونستمر بزيادة 10 فنحصل على الأرقام: 24، 34، 44 وبذلك نحصل على عينة منتظمة مؤلفة من 5 طلاب ولا ينصح باستخدام هذا النوع من العينات عندما يكون هناك تجانس في المجتمع أي عندما يكون هناك تكرار نمطي في الصفات والخصائص للأشياء أو الأفراد.

3- العينة الطبقية (Stratified Sample):

يتم في هذا النوع من العينة تقسيم المجتمع الإحصائي أولاً إلى مجموعات فرعية تسمى كل منها طبقة (Strate) ومن ثم تتم عملية المعاينة من كل طبقة، وتكون عادة جميع عناصر الطبقة الواحدة متجانسة فيما يتعلق بالخصائص موضوع الدراسة.

نقسم المجتمع إلى طبقات منفصلة ثم نختار من كل طبقة عدد من المفردات يتناسب مع حجم الطبقة ثم نجمع العينات الناتجة، ويلجأ الباحث إليها عندما يكون المجتمع الأصلي مكوناً من مجتمعات جزئية أو طبقات غير متداخلة (ذكور، إناث)، (متزوجون، مطلقون، أرامل، عزاب) وهكذا. ويتم اختيار عدد مفردات كل طبقة حسب العلاقة:

(حجم الطبقة/ حجم المجتمع) × حجم العينة

مثال:

عدد طلاب إحدى شعب الإحصاء (60)، منهم (40) طالباً و(20) طالبة.

وأردنا اختيار عينة حجمها (9) من هذه الشعبة على أن تكون من الطلاب والطالبات فما عدد الطلاب وما عدد الطالبات التي تتشكل منها العينة؟

الحل:

عدد الطلاب $= \frac{40}{60} \times 9 = \frac{360}{60} = 6$ طلاب.

عدد الطالبات $= \frac{20}{60} \times 9 = \frac{180}{60} = 3$ طالبات.

فبالتالي تتكون العينة المطلوبة من (6) طلاب + (3) طالبات وحجمها (9).

مثال:

إذا أردنا أخذ عينة مكونة من (20) طالباً من (3) كليات (الاقتصاد، العلوم الأساسية، الآداب) وعلمنا أن 50٪ من طلبة الجامعة من كلية الاقتصاد، 20٪ من كلية العلوم الأساسية، 30٪ من كلية الآداب) .

الحل:

حجم العينة المأخوذة من كلية الاقتصاد= $\frac{50}{100} \times 20 = 10$ طلاب.

وحجم العينة المأخوذة من كلية العلوم الأساسية = $\frac{20}{100} \times 20 = 4$ طلاب.

وحجم العينة المأخوذة من كلية الآداب = $\frac{30}{100} \times 20 = 6$ طلاب.

لاحظ أن حجم العينة المطلوبة = 10 + 4 + 6 =20 طالباً

4- العينة العنقودية متعددة المراحل (Multi - stage Cluster Sample):

تعتبر المعاينة العنقودية أحد الآليات التي يمكن استخدامها لاختيار العينات من خلال تقسيم المجتمع إلى مجموعات أو عناقيد (Clusters) ويتم اختيار عينات

عشوائية من هذه العناقيد بحيث أن جميع المشاهدات في العناقيد المختارة تكون جزءا من العينة.

ويمكن أن يتم الاختيار على عدة مراحل باستخدام طرق المعاينة الثلاث سابقة الذكر. فإذا تم الاختيار على مرحلتين فإنها تسمى ثنائية المراحل وإذا تم على ثلاث مراحل فإنها تسمى ثلاثية المراحل وهكذا.

مثال:

إذا أردنا اختيار لجنة مكونة من (20) طالباً من كليتين من جامعة بها (6) كليات وكان 40٪ من الطلبة من كلية الاقتصاد، 15٪ من كلية الصيدلة، 5٪ من كلية الحقوق، 10٪ من كلية العلوم، 18٪ من كلية الآداب، 12٪ من كلية الهندسة.

الحل:

المرحلة الأولى نختار كليتين عشوائياً من بين الكليات الست، ولنفرض أننا حصلنا على كليتي الاقتصاد والعلوم. وفي المرحلة الثانية نختار العينة المطلوبة بطريقة العينة الطبقية ثم نجري معاينة طبقية بحيث تتكون العينة المطلوبة من هاتين الكليتين فقط.

فيكون حجم العينة المختارة من كلية الاقتصاد = $\frac{80}{100} \times 20 = 16$ طالباً

وحجم العينة المختارة من كلية العلوم = $\frac{20}{100} \times 20 = 4$ طلاب

لاحظ أن حجم العينة المطلوبة = $16 + 4 = 20$ طالباً

وتجدر الإشارة إلى أن هناك عينات غير احتمالية تجمع البيانات من هذه العينات لغايات وصفية حيث لا يستخدم الإحصاء الاستدلالي استناداً إلى تلك العينات وبالتالي فإنه لا يمكننا تعميم النتائج من العينة إلى المجتمع ومن الأمثلة عليها. عينة الصدفة والعينة العقدية والعينة الحصصية وغيرها.

مقاييس النزعة المركزية.

تمهيد

تعاريف

المتوسط الحسابي

المنوال

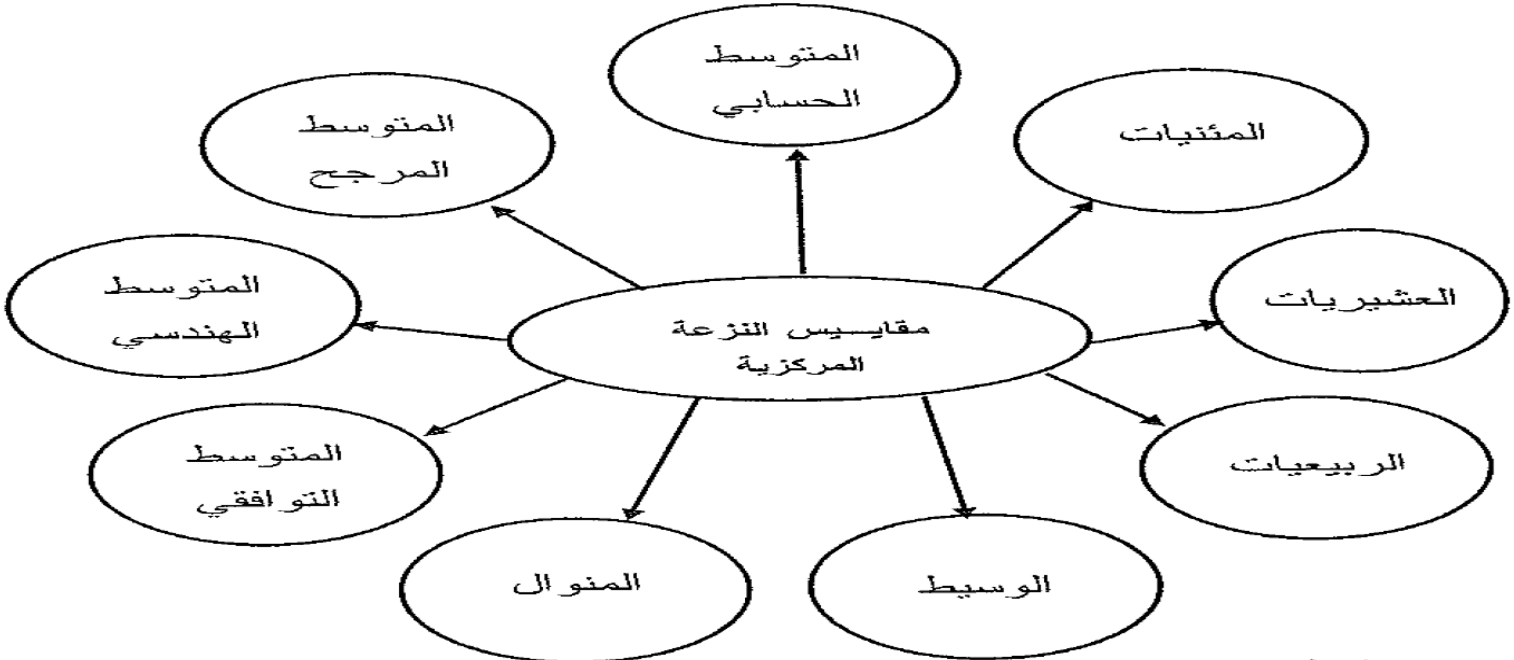
الوسيط

تمارين

مقاييس النزعة المركزية

تمهيـــــد

تعود فكرة مقاييس النزعة المركزية إلى الباحث الإنجليزي فرانسيس جالتون (1822-1911). وفيها يستخدم الباحثون العديد من المقاييس، لكن أكثرها استعمالا هي: المتوسط، الوسيط والمنوال لكونها تساعد على تلخيص البيانات ووصفها عن طريق التعرف على مركزها، فهي تقدم معلومات حول الوضعية العامة للتوزيع؛ فعندما نقوم بدراسة ظاهرة من الظواهر، قد نجد أن قيم تلك الظاهرة تميل نحو التجمع أو التمركز حول قيمة معينة، ومن ثم يطلق على هذا "الميل" اسم النزعة المركزية. فمقاييس هذه النزعة هي عبارة عن قيم مثلى تقترب منها معظم البيانات الإحصائية أو تتمركز حولها أو بالقرب منها، وهكذا فإن قيمة أي مقياس من مقاييس هذه النزعة تشير إلى نقطة تمركز مجموعة من البيانات، وتصف نقطة تجمع المشاهدات. وقبل عرض المقاييس التي تخدم هذا الفصل، نقدم شكلا توضيحيا يلخص مقاييس النزعة المركزية.



تعاريف:

وفيما يلي نبين في الجدول الآتي تعاريف مقاييس النزعة المركزية ومميزاتها.

المتوسط	مميزاته
– المتوسط الحسابي لقيم متغير ما، هو مجموع قيم ذلك المتغير مقسوما على عدد هذه القيم، فهو معلومة رقمية تتجمع حولها سلسلة من القيم، يمكن من خلالها الحكم على بقية قيم المجموعة.	– يعتمد في حسابه على جميع القيم. – هو نقطة اتزان المشاهدات. – أقل مقاييس النزعة المركزية تأثرا بالتقلبات العينية. – يتأثر بالقيم المتطرفة، ولذا فهو لا يصلح للتوزيعات الملتوية.

- في حالة الجداول التكرارية يكون المتوسط الحسابي هو مجموع حواصل ضرب مراكز الفئات في تكراراتها مقسوم على مجموع التكرارات.	- لا يصلح في حالة الفئات المفتوحة لعدم وجود مركز فئة. - لا يمكن حساب متوسط المتغيرات النوعية.

المنوال	مميزاته
1. هو القيمة الأكثر تكرارا في مجموعة من البيانات. 2. في حالة الجداول التكرارية يكون المنوال هو مركز الفئة المنوالية.	1. غير ثابت. 2. يتأثر بطول الفئة. 3. يمكن استعماله عندما يكون مستوى القياس اسمي.

الوسيط	مميزاته
1. وسيط مجموعة من الأرقام مرتبة حسب قيمها في ترتيب تصاعدي أو تنازلي، هو القيمة التي تتوسط البيانات التي تقع في المنتصف أو الوسط الحسابي للقيمتين اللتين تتوسطان البيانات أو تقع في منتصف البيانات، أي تقسم البيانات إلى قسمين متساويين. 2. في الجداول التكرارية يعبر الوسيط عن القيمة التي تكرارها تراكمي (ن/2) (ن) عدد المشاهدات (أو القيم).	1. لا يتأثر بالقيم المتطرفة. 2. يمكن استخدامه في التوزيعات الملتوية. 3. يفضل في استخدام الفئات المفتوحة. 4. يأتي في المرتبة الثانية في ما يخص تأثره بالقيم المتطرفة بعد المتوسط. 5. من خصائصه أنه يأخذ بعين الاعتبار رتب القيم، ويستعمل في البيانات الرتبة بشكل خاص. 6. لإيجاده يجب ترتيب القيم تصاعديا أو تنازليا. 7. إذا كانت البيانات رتبية يفضل حساب الوسيط.

المتوسط الحسابي

استعمالاته:

1- حساب المتوسط الحسابي لعدد من القيم:

يرمز له عادة بحرف:($\overline{X}$) ولحساب متوسط مجموعة من القيم ($x_1 + x_2 ... x_n$)، نقوم بحساب مجموعها ونقسمه على عددها، أي أنه ينبغي تطبيق الصيغة التالية:

$$\boxed{\overline{X} = \frac{\sum x_1 + x_2 ... x_n}{n}}$$

حيث أن:

$\overline{X}$ = المتوسط.

$\sum x$ = مجموع القيم.

n = عدد القيم.

مثال: تحصل تلميذ على العلامات التالية: 12. 14. 9. 10. 5. 13. 15 . 14.

$$\overline{X} = \frac{14 + 15 + 10 + 5 + 13 + 9 + 14 + 12}{8}$$

$$\overline{X} = \frac{92}{8}$$

$$\boxed{\overline{X} = 11.5}$$

2- حساب المتوسط الحسابي لبيانات مجمعة في جدول تكراري :

لحساب المتوسط الحسابي من جدول تكراري نطبق المعادلة التالية:

$$\boxed{\overline{X} = \frac{\sum (xifi)}{\sum fi}}$$

حيث أن:

$\overline{X}$ = المتوسط الحسابي.

xi = القيمة المشاهدة.

fi = تكرار القيمة.

$\sum fi$ = حجم العينة أو مجموع كل التكرارات.

مثال : في دراسة لحساب عدد الأهداف المسجلة من قبل 10 لاعبي كرة القدم أراد مختص أن يحسب متوسط الأهداف، فجمع البيانات الآتية:

عدد الأهداف (xi)	التكرارات (fi)	(fi xi)
2	1	2
1	1	1
4	1	4
5	4	20
7	2	14
8	1	8
	10	47

$$\overline{X} = \frac{49}{10} = 4.9$$

متوسط الأهداف المسجلة من قبل العشرة لاعبين = 4.9

3- حساب المتوسط الحسابي لبيانات مبوبة:

عندما يزيد عدد البيانات عن 30 نبوب البيانات في فئات ونطبق المعادلة الآتية:

$$\boxed{\overline{X} = \frac{\sum f \times x}{\sum f}}$$

حيث أن :

f = التكرارات

x= مراكز الفئات

(حساب مركز الفئة= الحد الأدنى للفئة + الحد الأعلى لنفس الفئة /2).

$\sum f$ = مجموع التكرارات.

الخطوات:

1. وضع عمود تحدد فيه مراكز الفئات (x).
2. وضع عمود تحدد فيع التكرارات.
3. وضع عمود نضع فيه نتيجة ضرب مركز كل فئة في التكرار الذي يقابلها. (f).
4. وضع عمود نجمع فيه حاصل ضرب مراكز الفئات في تكراراتها$(f \times x)$
5. نقسم الحاصل على مجموع التكرارات.

مثال: يوضح الجدول الآتي علامات العدوانية لدى 100 فرد في أحد مراكز إعادة التربية:

العدوانية	التكرارت f	مراكز الفئات x	$(f \times x)$
62-60	5	61	305
65-63	8	64	512
68-66	42	67	2814
71-69	27	70	1890
74-72	8	73	584
77-75	10	76	760
المجموع=100			6865

المتوسط الحسابي للعدوانية عند الفئة المدروسة:

$$\overline{X} = \frac{6865}{100} = 68.65$$

المنوال:

يمكن حساب المنوال من البيانات الخام ومن البيانات التكرارية ومن البيانات المبوبة في فئات:

1. حساب المنوال في حالة البيانات الخام وهي ثلاثة أنواع:

- ✓ السلسلة الرقمية غير المنوالية مثل: (10. 5. 15.20. 25. 50) ففي هذه السلسة لا يوجد منوال مادام أن أية قيمة منها لم تتكرر، فكل واحدة منها تختلف عن الأخرى.
- ✓ السلاسل الرقمية أحادية المنوال مثل: (5. 25. 10. 25. 50) المنوال في هذه السلسلة هو القيمة (25) لأنها تكرر مرة واحدة.
- ✓ السلاسل الرقمية ثنائية المنوال مثل: (5. 5. 25. 25. 50) في هذه السلسلة يوجد منوالان هما: (5 و25).
- ✓ السلاسل الرقمية متعددة المنوال مثل: (5. 25. 5.50. 50. 25. 75. 80) في هذه السلسلة يوجد أكثر من منوالين وهم: (5.25.50).

مثال1: – المجموعة 22,5,7,9,9,9,10,10,11,12,18 لها منوال واحد وهو 9.

مثال2: – المجموعة 3,5,8,10,15,16 ليس لها منوال

مثال3: – المجموعة 2,3,4,4,4,5,5,7,7,7,9 لها منوالان وهما 7,4

وتسمى مجموعة ذات منوالين Bimodal التوزيع الذي له منوال وحيد يسمى وحيد المنوال Unimodal

يمكن أن توجد سلسلة من القيم فيها أكثر من منوالين وتسمى سلسلة متعددة المناويل Multimodal.

مثال: اختيرت أربع مجموعات من طلاب بعض التخصصات في كلية العلوم الاجتماعية /علم الاجتماع (المج1)، علم لنفس المدرسي(المج2)، علوم التربية(المج3)، علم النفس العيادي(المج4)/، وتم رصد درجات

هؤلاء الطلاب في وحدة الإحصاء التطبيقي وكانت النتائج كالآتي:

المجموعة 1	77	65	77	77	منوالين هما 77-77
المجموعة 2	85	77	65	93	منوالين 65-65
المجموعة 3	65	76	88	65	منوالين 65-65
المجموعة 4	73	69	69	73	منوالان 69-69

2. حساب منوال البيانات التكرارية (جدول تكراري):

البيانات أدناه توضح توزيع عينة من العمال حسب وضعيتهم المدنية:

الوضعية المدنية	عدد العمال (التكرارات)
متزوج	10
مطلق	7
أرمل	3
أعزب	25
المجموع	45

الحل: المنوال = أعزب (لأنها الحالة الاجتماعية المقابلة لأكبر تكرار وهو 25).

3. حساب منوال بيانات مبوبة في فئات.

لحساب المنوال لهذا النوع من البيانات تطبق الصيغة الآتية:

$$Mode = L + \left\{ \frac{d_1}{d_{1} + d_2} \right\} i$$

حيث أن:

L = الحد الأدنى لفئة المنوال (الفئة المناظرة لأكبر تكرار) .

d_1 = الفرق الأول = (تكرار فئة المنوال – التكرار السابق لها).

d_2 = الفرق الثاني = (تكرار فئة المنوال – التكرار اللاحق لها).

i = طول فئة المنوال.

مثال: الجدول التالي يبين فئات درجات مجموعة من التلاميذ.

فئات الدرجات	عدد التلاميذ
15- 5	2
25- 15	6
35- 25	10
45- 35	7
55-45	3
55- 65	2
المجموع	30

الفئة المنوالية: 25 -35 (يقابلها أكبر تكرار). الحد الأدنى للفئة المنوالية = 25.

طول الفئة = 10. تكرار الفئة المنوالية = 10. تكرار قبل الفئة المنوالية = 6.

التكرار اللاحق لتكرار الفئة المنوالية = 7.

$$Mode = L + \left\{\frac{d_1}{d_{1+}d_2}\right\}i =$$

$$= 25 + \left\{\frac{4}{4+3}\right\}10$$

$$= 25 \times 5.7$$

$$\boxed{\text{المنوال} = 30.71}$$

يلاحظ أن قيمة المنوال تقع في الفئة المنوالية 25 -35.

الوسيط

1 – حساب الوسيط من قيم مفردة (غير مكررة):

✓ عندما يكون عدد القيم فرديا نتبع الخطوات الآتية:

- نرتب القيم تصاعديا أو تنازليا.
- نطبق الصيغة الرياضية الآتية لحساب رتبة الوسيط:

$$\boxed{Me = \frac{n + 1}{2}}$$

حيث أن:

Me = الوسيط.

n = عدد القيم.

مثال: (7. 4 .3 .6 .5) نرتب هذه السلسلة تصاعديا مما يعطينا (3. 4. 5. 6. 7)

رتبة الوسيط $= 3 = \frac{6}{2} = \frac{5+1}{2} = Me$ الوسيط إذن هو القيمة التي تحتل الرتبة 3 وقيمته هي الدرجة (5).

✓ عندما يكون عدد القيم زوجيا نطبق نفس المعادلة السابقة:

لدينا القيم التالية: (15.10.5 .25 .50 .55) مرتبة تصاعديا

رتبة الوسيط: $3.5 = \frac{7}{2} = \frac{6 + 1}{2} = Me$

قيمة الوسيط توجد بين القيمتين الثالثة والرابعة، إذن ستكون بين 15 (القيمة الثالثة) و25 (القيمة الرابعة).

$$= \frac{15 + 25}{2} = \frac{40}{2}$$

$$\boxed{Me = 20}$$

القيمة التي تقسم مجموعة القيم (10.5 .15 .25 .50 .55) = 20

2- وسيط قيم مجمعة في جدول تكراري :

عندما تكون القيم مجمعة في جدول تكراري، فإن حساب الوسيط يتم بتطبيق المعادلة التالية:

$$Me = L + \left\{ \frac{\frac{n}{2} - fc}{f} \right\} i$$

حيث أن :

L = الحد الأدنى للفئة الوسيطية (هي الفئة التي يقابلها أكبر تكرار).

n = عدد المشاهدات.

fc = مجموع التكرارات التراكمية للفئات التي تسبق الفئة الوسيطية.

f = التكرارات المقابلة للفئة الوسيطية.

i = طول الفئة.

مثال: يبين الجدول التالي توزيع علامات 50 طالبا في مادة مدخل إلى علم النفس:

العلامات	مراكز الفئات	التكرارات	التكرار المتجمع الصاعد	التكرار المتجمع النسبي
10-1	5	5	5	10%
20-10	15	10	15	%30
30-20	25	20	35	%70
40-30	35	11	46	%92
50-40	45	4	50	%100
		مج n=50		

الوسيط يوجد في الموقع الذي توجد فيه نسبة 50% وهي تقع في الفئة 20-30.

الحد الأدنى للفئة الوسيطية إذن = 20 مما يدل على أن قيمة الوسيط لا يمكن أن

تقل عن 20. التكرار المتجمع الصاعد السابق للفئة الوسيطية = 15 في حين أن التكرار (f) المقابل للفئة الوسيطية = 20، أما طول الفئة الوسيطية (i) = 10 لأنه نتيجة لـــ: 20−30=10.

الوسيط إذن يساوي:

$$Me = L + \left\{ \frac{\frac{n}{2} - fc}{f} \right\} i$$

$$Me = 20 + \left\{ \frac{\frac{50}{2} - 15}{20} \right\} 10 = 20 + \left\{ \frac{25 - 15}{20} \right\} 10$$

$$= 20 + \left\{ \frac{10}{20} \right\} 10 = 20 + 0.5(10) = 20 + 5$$

الوسيط = $\boxed{Me = 25}$

مقاييس التشتت

تمهيد

المدى

التباين

الانحراف المعياري

معامل الاختلاف

مقاييس التشتت

تمهيد :

تناولنا في الفصل السابق مقاييس النزعة المركزية والتي تعبر عن المستوى العام للظاهرة محل البحث. ولكن ترى هل هذا يعتبر كافياً لوصف البيانات وتحليلها كمياً لنصل بالتالي إلى فهم أكثر وضوحاً للظاهرة؟ للإجابة على هذا السؤال نسوق المثال التالي:

مثال: نفترض أن لدينا فئتين من العمال اليدويين تعيشان في منطقتين مختلفتين وكان أجر كل فرد من المجموعتين على الساعة (بالدينار) كما يلي:

المجموعة الأولى:

78	73	76	74	75	71	75	70

المجموعة الثانية:

93	65	70	29	100	80	56	99

وبحساب المتوسط الحسابي للفئتين المذكورتين في كل من المجموعتين

فإن المتوسط الحسابي لأجر المجموعة الأولى:

$$\overline{X}_1 = \frac{70+75+71+75+74+76+73+78}{8}$$

$$\overline{X}_1 = \frac{592}{8} = 74$$

أي أن المتوسط الحسابي لأجر هذه المجموعة = 74 دينارا.

وكذلك بالنسبة للمجموعة الثانية فإن المتوسط الحسابي لأجرها بالساعة هو:

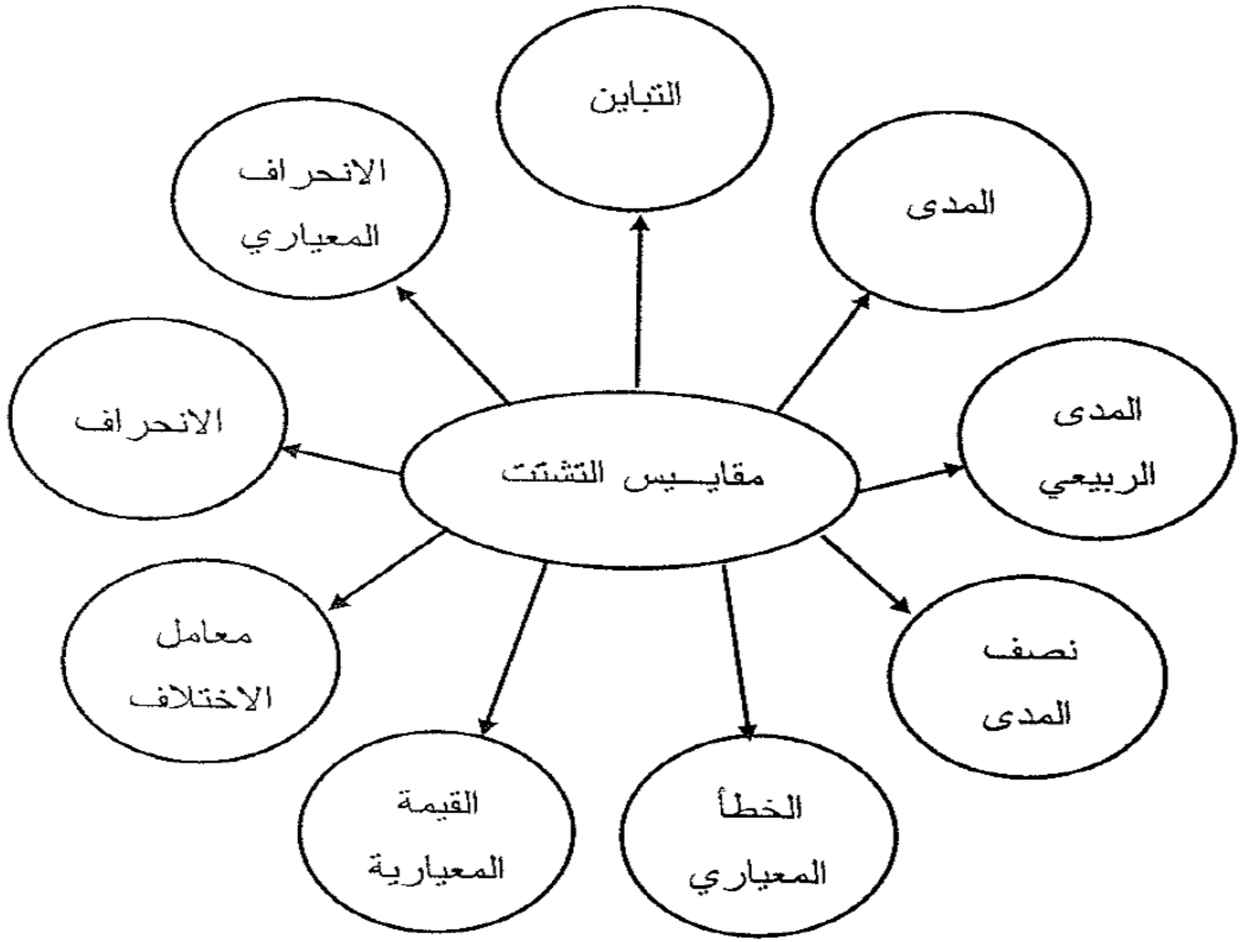
$$\overline{X}_2 = \frac{99+56+80+100+29+70+65+93}{8}$$

$$\overline{X}_2 = \frac{592}{8} = 74$$

والمتوسط الحسابي لأجرها هو أيضاً 74 دينارا. ومعنى هذا أن متوسط أجور الأفراد في المجموعتين = 74 دينارا.

لكن، بإمعان النظر في أجور المجموعة الأولى، نجد أنها متجانسة إلى حدٍ كبير؛ أي أنها قريبة جداً من بعضها أو من المتوسط الحسابي، وهنا يتضح أن تشتت أجور هذه المجموعة قليل أو صغير (أجور متجانسة أو قريبة من التجانس) بينما أجور المجموعة الثانية غير متجانسة، فهي بعيدة عن بعضها أو عن المتوسط الحسابي بشكل كبير، وهنا يتضح أن تشتت أجور المجموعة الثانية كبير وبالتالي فهي أقل تجانساً (أو أكثر تشتتاً) من المجموعة الأولى. فالفرق بين أكبر أجر وأصغر أجر في المجموعة الأولى= 78- 70 =8 دينارا فقط، بينما الفرق بين أكبر أجر وأصغر أجر في المجموعة الثانية: 100- 29=71 دينارا).

انطلاقا من هاتين الملاحظتين، نستطيع أن نستنتج أن مقاييس النزعة المركزية غير كافية لوصف البيانات أو تحليلها كمياً، فها هو المتوسط واحد في المجموعتين ورغم ذلك فإن البيانات تختلف تماماً في مدى تشتتها و تجانسها. من هنا تبرز الحاجة إلى مقاييس أخرى ليقاس مدى تشتت البيانات لتعطي للباحث صورة أكثر وضوحاً وصدقاً للظاهرة محل الدراسة. وقبل التعرض بالتفصيل لمقاييس التشتت، نقدم في ما يلي شكلا توضيحيا لأهم مقاييس التشتت:



يتضح من هذا الشكل أن مقاييس التشتت عديدة، وسنكتفي بتقديم أكثرها استعمالا وهي: المدى، والتباين والانحراف المعياري، ومعامل اختلاف.

المدى

يعرف المدى بأنه الفرق بين أكبر قيمة وأصغر قيمة في سلسلة من القيم.

خصائص المدى:

- ✓ يعتمد المدى في حسابه على قيمتين فقط (أكبر قيمة وأصغر قيمة).
- ✓ لا يقيس تشتت البيانات عن متوسطها، فهو لا يشير إلى متوسط البيانات.
- ✓ قليلا ما يلجأ الباحثون إلى استعمال المدى نظرا لبعض المساوئ التي يتميز بها وهي:

- إذا كان عدد أفراد العينة كبيرا، فهو يعطي فكرة غامضة عن تشتت قيم الظاهرة.
- لا يأخذ بعين الاعتبار إلا قيمتين فقط. للتعبير عن بقية القيم.
- تزداد عدم دقته كلما ازداد فيه التطرف والشذوذ عن القيم العادية.

حساب المدى:

يحسب المدى لسلسلة من القيم، ولبيانات مبوبة:

1 – حساب المدى لسلسلة من القيم:

مثال: في أحد اختبارات المهارة اللغوية تحصلت مجموعة من التلاميذ على الدرجات التالية:

33	58	43	51	70	66	96	32	55	63

نلاحظ أن أكبر قيمة هي 96 وأدنى قيمة هي 32، فالمدى يساوي:

$96 - 32 = 64$. إذا كانت لدينا القيم الآتية: 2،3، 8، 11، 6، 7، يكون المدى هو: $11-2 = 9$. يمكن أن نوضح أكثر بتقديم المثال التالي: إذا قسمنا مثلا علامات فوجين من الطلبة في اختبار الإحصاء وحصلنا على السلسلتين التاليتين:

علامات الفوج الأول: 7، 6، 4، 5، 11، 9، 8، 10

علامات الفوج الثاني: 8، 6، 5، 9، 17، 13، 10، 8

نستطيع أن نلاحظ أن الحد الأدنى لدرجات الفوج الأول هو 4 درجات وأن الحد الأعلى لنفس الفوج هو 11 وحسب هذين الحدين فإن المدى للفوج الأول هو: 11−4 = 7. أما الفوج الثاني فإن مداه بإتباع نفس الطريقة هو: 17 − 5 = 12. في ضوء هذه المقارنة يتضح أن درجات أفراد المجموعة الثانية أكثر تشتتا من درجات أفراد المجموعة الأولى، وبعبارة أخرى يمكننا اعتبار درجات المجموعة الأولى أقرب إلى التجانس من درجات المجموعة الثانية.

2− حساب المدى لبيانات مبوبة:

يحسب المدى في هذه الحالة بالطريقة التالية:

الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى +1.

مثال:

الفئات	10 -5	15 -10	20 -15	25 -20	30 -25

المدى = 30− 5 + 1 = 26

يتضح من الأمثلة السابقة أن حساب المدى عملية بسطيه لأنه يعتمد على طرح أدنى قيمة من أكبر قيمة ضمن سلسلة من القيم.

التباين

هو أحد مقاييس التشتت، ويمكن التفكير فيه كمقياس للمسافة، حيث تقاس المسافة ببعد القيمة عن المتوسط الحسابي، وكلما كانت قيمة التباين كبيرة كان التوزيع أكثر تبعثرا وأقل تجانسا. ومن مزاياه أنه يأخذ جميع القيم في الاعتبار لقياس تشتت القيم حول المتوسط الحسابي. إضافة إلى ذلك، يصلح التباين لتحليل وقياس الفروق الجماعية بين الأنواع المختلفة للتوزيعات التكرارية، كحساب الفروق بين مستويات تحصيل مجموعة من الطلبة ومجموعة من الطالبات ويسمى هذا النوع من التحليل بتحليل التباين.

إن الفكرة الأساسية لهذا المقياس هي حساب انحرافات جميع القيم عن متوسطها الحسابي (أي أنه يحسب الفرق بين كل قيمة والمتوسط الحسابي) ودائما يكون مجموع هذه الانحرافات مساويا للصفر، ويكون الحل هنا إهمال الإشارات السالبة أو تربيع هذه الانحرافات. لكن إهمال الإشارات السالبة ليس له مبرر رياضي، ومن ثم يكون الحل الأمثل هو تربيع هذه الانحرافات، ثم نحسب متوسط الانحرافات المربعة لنحصل على قيمة التباين، لذلك يعرف التباين كما يلي:

التباين = متوسط مربعات انحرافات القيم عن متوسطها الحسابي.

يرمز لتباين العينة بحرف (S^2) ويرمز له بحرف سيغما (σ^2) عندما نحسبه لمجتمع.

يحسب التباين بصيغتين:

أ. عندما نحسب تباين المجتمع نطبق الصيغة التالية:

$$\boxed{\sigma^2 = \frac{\sum\left(x - \bar{x}\right)^2}{n}}$$

σ^2 = تباين المجتمع.

x = الدرجات.

$\bar{x}$ = متوسط الدرجات.

n = عدد الأفراد.

ب. عندما نحسب تباين العينة نطبق المعادلة التالية:

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

حسب التباين من درجات العينات ومن فئات:

1- حساب التباين من درجات العينات:

ثال: فيما يلي درجات عينة من التلاميذ: (5،10،15،12،14).

$\sum (x-x)^2$	$x-\bar{x}$	X	الأفراد
$(-6.2)^2 = 38.44$	$5-11.2=-6.2$	5	1
$(-1.2)^2 = 1.44$	$10-11.2=-1.2$	10	2
$(3.8)^2 = 14.44$	$15-11.2=3.8$	15	3
$(0.8)^2 = 0.64$	$12-11.2=0.8$	12	4
$(2.8)^2 = 7.84$	$14-11.2=2.8$	14	5
$\sum = 62.8$		$X = \frac{56}{5} = 11.2$	5 =n

وبما أن n= 5

التباين =

$$S^2 = \frac{62.8}{5-1} = 15.7$$

2- حساب التباين من فئات:

لحساب التباين من جدول فئات نطبق المعادلة التالية:

$$\boxed{S^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n-1}}$$

حيث أن:

= التكرارات.

x= مراكز الفئات.

n= عدد الأفراد.

مثال: إليك توزيع علامات 15 فردا من ذوي الاحتياجات، احسب التباين:

$F.x^2$	x^2	fx	مراكز الفئات x	التكرارات f	الفئات
144	144	12	12	1	14-10
867	289	51	17	3	19 -15
2420	484	110	22	5	24-20
2916	729	108	27	4	29-25
2048	1024	64	32	2	34-30
8395		345		مج = 15	

$$s^2 = \frac{8395 - \frac{119025}{15}}{14} = 32.86$$

الانحراف المعياري

هو مقياس آخر للتشتت وهو الجذر التربيعي للتباين، ويفضل استخدامه بدلا من التباين لأن وحدة القياس فيه مساوية لوحدة القياس للبيانات الأصلية، وبالتالي يمكن التفكير فيه كمتوسط للمسافات بين القيم.

يرمز لهذا المقياس بحرف (S) عندما نحسبه لعينة، ويرمز له بحرف سيغما (σ) عندما نحسبه لمجتمع. وهو يعد من أهم مقاييس التشتت وأكثرها استعمالا ويمكن حسابه من الدرجات الخام ومن الدرجات التكرارية، لكنه لا يصلح للمقارنة بين التوزيعات المختلفة. إنه قيمة حقيقية موجبة وهو يستعمل لتمثيل توزيع متغير عشوائي حقيقي حول متوسطه، لذلك فهو يعرف بكونه متوسط انحراف القيم عن متوسطها، ويمثل الجذر التربيعي للتباين.

خصائصه:

- من خصائص الانحراف المعياري أنه كلما صغرت قيمته دل ذلك على أن طبيعة البيانات متقاربة ومتراكمة حول المتوسط، وبالتالي التشتت قليل والعينة متجانسة.
- الانحراف المعياري من أكثر مقاييس التشتت تأثرا بالدرجات المتطرفة في التوزيع، لاعتماده المباشر على مربعات فروق هذه الدرجات عن المتوسط وهو لا يتأثر تأثرا كبيرا بالدرجات القريبة عن المتوسط، لأن القيمة العددية لمربعات فروق تلك الدرجات عن المتوسط صغيرة، لكنه يتأثر بالمتوسط نفسه لأنه الإطار الذي ينسب إليه فروقه ومربعاتها.
- لا يتأثر الانحراف المعياري بإضافة عدد ثابت لكل درجة من درجات التوزيع التكراري أو بحذف قيمة عددية ثابتة من كل درجة من درجات ذلك التوزيع. والسبب الذي من أجله يتحرر الانحراف المعياري من أثر تلك الإضافة أن الحذف يبدو واضحا عندما ندرك أن انحراف أي عدد عن أي عدد آخر لا يتأثر بالإضافة أو الحذف. أي أن الانحراف المعياري لا يتأثر بالإضافة أو الحذف .
- قيمة الانحراف المعياري تكون دائما موجبة.

- يفيد في مقارنة توزيعين مختلفين.
- يوضح كيفية اختلاف النتائج.
- يبين كيفية تشتت وانتشار النتائج.

استعمالاته:

- تساعد قيمة الانحراف المعياري على معرفة كيفية تجمع البيانات حول المتوسط، فعندما تتجمع البيانات في التوزيع تجمعا وثيقا وقريبا من قيمة المتوسط، فإن قيمة الانحراف المعياري تكون صغيرة، وعندما تنتشر البيانات بعيدا عن بعضها البعض أكثر (مع ظهور قيم حدية على جانبي قيمة المتوسط) فإن قيمة الانحراف المعياري تكون كبيرة، أي أن الانحراف المعياري يبين مدى تقارب البيانات وتباعدها في الظاهرة معينة عن متوسطها الحسابي.

يحسب الانحراف المعياري للعينة بالصيغة التالية:

$$\boxed{S = \sqrt{\sum \frac{\left(x - \overline{x}\right)^2}{n - 1}}}$$

حيث أن:

x = الدرجة

$\overline{x}$ = متوسط الدرجات

n = حجم العينة

أ- حساب الانحراف المعياري لمجموعة من القيم:

يجب تحقيق الخطوات التالية:

1- حساب متوسط المتغير.

2- حساب انحراف كل قيمة عن المتوسط، بطرح القيمة من المتوسط.

3- حساب مجموع المربعات.

4- تقسيم المجموع على عدد المشاهدات.

5- نحسب الجذر التربيعي للتباين.

لنأخذ مثال الدرجات الخام التالية: 4 ، 6 ، 9 ،10، 8 ، 7. المتوسط = 7.33

نقوم بحساب انحراف الدرجات عن المتوسط، كما هو مثبت في الجدول أدناه:

الدرجات X	الانحراف من المتوسط $(x-\bar{x})$	مربع انحراف عن المتوسط $(x-\bar{x})^2$
4	3.33- = 7.33- 4	10.08
6	1.33- = 7.33- 6	1.76
9	1.67 = 7.33- 9	2.78
10	2.67 = 7.33 -10	7.12
8	0.67 = 7.33- 8	0.44
7	0.33 = 7.33- 7	0.10
$\bar{X}$ =7.33 / $\sum$ 44		$\sum$ 22.28

الانحراف المعياري =

$$S=\sqrt{\frac{22.28}{6-1}}=2.11$$

ب – حساب الانحراف المعياري لبيانات مبوبة:

لإيجاد الانحراف المعياري لبيانات مبوبة لعينة نطبق الصيغة الآتية:

$$S=\sqrt{\frac{\sum\left(x_i-\bar{X}\right)^2\times f}{n-1}}$$

حيث أن:

x_i = مركز الفئة.

$\bar{X}$ = المتوسط الحسابي.

f = تكرار الفئة.

$n-1$ = حجم العينة – 1.

ملاحظة: نضع (−1) لأن العينة تكون دائما أصغر من المجتمع الأصل.

مثال: لنوجد الانحراف المعياري من درجات 50 طالبا في وحدة اللغة الإنجليزية:

الفئات	مراكز الفئات (xi)	التكرارات (f)
10- 1	5	5
20 – 10	15	10
30- 20	25	20
40- 30	35	11
50- 40	45	4
	$\overline{X} = 25$	= 50

$$(5-24)^2 \times 5 = 1805 + (10-24)^2 \times 10 = 1960 + (25-24)^2 \times 20 = 20 +$$
$$(35-24)^2 \times 11 = 1331 + (45-24)^2 \times 4 = 1764$$
$$= 1805 + 1960 + 20 + 1331 + 1764 = 6880$$

بتطبيق المعادلة الخاصة بحساب الانحراف المعياري للعينات:

$$= \sqrt{\frac{6880}{5-1}} = \sqrt{1720}$$

$$\boxed{S = 41.47}$$