كلية المامون الجامعة
قسم المالية والمصرفية

مبادئ الاحصاء

المحاضرة

2023

د سلسل صادق كنعان

1- مفهوم علم الاحصاء

1-1. المقـــــــدمة

يعد علم الاحصاء(statistics) ، اليوم، من اهم العلوم التي تتوقف عليها التنمية السياسية والاقتصادية والثقافية الخ...، وللإحصاء حصة اساسية من عمل الدول والمؤسسات والمنظمات السياسية والاقتصادية والاجتماعية، عالميا ودوليا ومحليا، وكثيرا ما يرتهن مصير مشاريع او قرارات كبرى بالنتائج التي يقدمها الاحصاء في مجال معين.

وبصورة عامة، فان افتقاد الجهد الاحصائي، في مجال من المجالات، يمنع من التأكد وتحصيل الضمان في استجابة اي مشروع للواقع، كما يحول دون تحديد مدى نجاحه او اخفاقه، ويجعل في الاقدام عليه شيئا من المخاطرة.

جاء في موسوعة للاند حول الاحصاء من الناحية الجوهرية انه مجموعة الوقائع التي يودي اليها اجتماع البشر في مجتمعات سياسية... لكن الكلمة عندنا سترتدي مفهوما اوسع، فنحن نعني بالإحصاء العلم الذي يكون موضوعه جمع وتنسيق وقائع كثيرة في كل صنف، بحيث يمكن الحصول على نسب عددية مستقلة استقلالا ملموسا عن المصادفة واستثناءاتها، وفي موسوعة المورد العربية يعرف بانه علم جمع وتصنيف وتعليل الوقائع او المعطيات الرقمية او العددية، يتخذ طريقة للتحليل في العلوم الدقيقة والعلوم الاجتماعية وفي المشروعات الاقتصادية على اختلافها. وهو يعنى، في آن معا، بوصف الوقائع وبالتنبؤ باحتمالات حدوث امر بعينه او حالة بعينها. وعلم الاحصاء علم حديث نشا في مطالع القرن العشرين، وتطور تطورا كبيرا بعد الحرب العالمية الثانية، وانما يعزى هذا التطور الكبير الى استحداث الحاسبات الالكترونية التي تتعامل مع كميات من الارقام ضخمة تعاملا سريعا. ومن تعريف علم الاحصاء انه منهج يتعاطى بالدرجة الاولى مع ظواهر رقمية وعددية معينة، ثم يقوم بتصنيفها وتحويلها الى نسب عددية خاصة، فيستطيع بالتالي تقديم وصف ميداني مرقم واكثر دقة للواقع، ويرفق ذلك الوصف بتقديم تصور علمي للعلل والاسباب التي ولدت الظاهرة المدروسة.

2-1. الهدف من دراسة مادة الاحصاء :

تهدف دراسة مادة الاحصاء الى:

1) تعريف الطالب بطبيعة علم الاحصاء وأهميته ومجالات استخدامه.
2) تعريف الطالب بأساليب جمع البيانات وطرق اخذ العينات ومن ثم طرق عرضها بالطرق البيانية.
3) تعريف الطالب باتباع الاساليب الاحصائية لغرض الوصول الى النتائج الدقيقة بأقصر طريقة واقل كلفة.

4) للإحصاء دور مهم واساسي في التخطيط للمشاريع سواء اكانت مشاريع فردية ام مشاريع تخص المجتمع استخدام الاساليب الاحصائية تساعد الطالب مستقبلا على اتخاذ قراراته سواء اكانت قرارات تجارية خاصة بالشراء وبيع المنتجات او تحديد اجور عمال .ام كانت قرارات صناعية او زراعية وغيرها .

5) توضيح للطالب كيف ان الاحصاء تغير من صورته القديمة الموجودة في اذهان الناس على انه علم العد وجمع البيانات وعرضها الى اعتباره الان علم يعمل على استخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتحليلها وعرضها لاستخلاص النتائج واتخاذ القرار المناسب على ضوء ذلك .

3-1. طبيعة علم الاحصاء. Nature Of Statistics

كلمة (الاحصاء) في الماضي كانت تهدف الى العد والحصر حتى سمي الاحصاء بعلم العد (The Science of canting) . اما الاحصاء الأن فقد تطور كثيرا وخاصة في القرن العشرين واصبح علما مستقلا له اهميته كوسيلة واداة في البحث العلمي لجميع العلوم .

1-3-1. تعريف علم الحصاء : -

هناك تعاريف عديدة للإحصاء اختلفت وتباينت من حيث المضمون والشمول باختلاف مراحل تطوير هذا العلم والفوائد المتوخاة منه .

فقد عرف بأنه الطريقة العلمية التي تختص بجمع البيانات والحقائق بشكل يسهل عملية تحليلها وتفسيرها ومن ثم استخلاص النتائج واتخاذ القرار على ضوء ذلك ، وهناك من عرفه بانه العلم الذي يعمل على استخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول منها الى استنتاجات وقرارات مناسبة.

ويمكن تقسيم علم الاحصاء الى قسمين هما :

1 - الاحصاء الوصفي Descriptive statistics : - ويتضمن الطرق الاحصائية المستخدمة في جمع البيانات والمعلومات عن ظاهرة معينة او مجموعة ظواهر وكيفية تنظيم وتصنيف وتبويب هذه البيانات مع امكانية عرضها في جداول ورسوم بيانية وحساب بعض المؤشرات الاحصائية.

2 - الاحصاء الاستنتاجي او الاستدلاليStatistical inference :- هو الشطر الاخر من علم الاحصاء الذي يهتم بموضوع التقديرات واختيار الفرضيات.

4-1 . اهمية علم الاحصاء وعلاقته بالعلوم الاخرى ومجالات تطبيقه

يعتبر علم الاحصاء احد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع المعلومات والبيانات اللازمة للبحث العلمي وتحليل هذه البيانات والمعلومات بهدف الوصول الى النتائج التي يهدف اليها البحث العلمي . وللإحصاء دورا بارزا في **التخطيط** ووضع الخطط المستقبلية عن طريق التنبؤ بالنتائج ففي قصة سيدنا يوسف عليه السلام مثال عظيم لدور الاحصاء في التخطيط فيبين ان هناك سنوات عجاف يقل فيها المحصول وسنوات سمان يزيد فيها المحصول ويبين انه يجب الاحتفاظ لسني القحط بادخار جزء من انتاج سني الرخاء . وفي مجال **الاقتصاد والاجتماع** يبرز دور الاحصاء في بحوث السكان متمثلا في تعدادات السكان فالتخطيط السليم لتنمية اقتصادية واجتماعية لا ينفصل ولا يمكن أن يتم بدون الدراسات الاحصائية للسكان ، فكيف نقرر اقامة مصانع ونحن لانعرف حجم قوة العمل المتوافرة والتي ستتوفر خلال فترة مقبلة وعلى اي اساس نقيم سياسة للإسكان ونحن لانعرف حجم قوة العمل المتوافرة والتي ستتوفر خلال فترة مقبلة وعلى اي اساس نقيم سياسة للسكان ونحن لانعرف معدلات الزواج والطلاق ..... وهكذا وفي مجال **الزراعة** يأتي دور الاحصاء في ان العلوم الزراعية تبدأ بالملاحظة وجمع بيانات عن الطبيعة في الحقل او المزرعة ثم يلي ذلك الدراسات العملية ويفيد الاحصاء في تنظيم وترتيب عملية الملاحظة والمشاهدة وجمع البيانات وتحليلها واستخلاص النتائج ولا يمكن أن يكون ذلك بغير دراسة كاملة بأساليب الاحصاء. وللإحصاء ايضا اهمية في مجال **الصناعة** من خلال استخدام النظرية الاحصائية في الانتاج الحربي وفي مجالات صناعات الفحم والحديد والغزل والمواد الكهربائية كما ان للإحصاء دور فعال في **مجال الطب والصحة العامة** في معرفة عدد المواليد وعدد الوفيات حيث تعتبر مؤشرات للمستوى الصحي العام ومؤشر لمدى تقدم البلد او تخلفه وأصبح للإحصاء أهمية كبرى في دراسة وتحليل العلاقات بين الامراض المختلفة وطرق العلاج واستخدام نظريات العروض الاحصائية اصبح الاساس في عمل شركات انتاج العقاقير والادوية . وعليه فأن الاحصاء بحد ذاته وسيلة وليس غاية فذاك يعني امكانية استخدامه اينما وجد في البحث العلمي.

5-1. الطريقة الاحصائية في البحث العلمي .

استخدام الاسلوب الاحصائي في البحث العلمي يعني توفير البيانات والمعلومات عن الظاهرة المطلوب دراستها في ذلك البحث وهذا يعني ان امكانية تطبيق الطريقة الاحصائية مرهونا بإمكانية التعبير عن هذه الظاهرة أو تلك تعبيرا كميا (رقميا) .وتمتاز الطريقة الاحصائية بكونها تهيئ اسلوبا موضوعيا محايدا للبحث له قواعده واصوله التي يجب أن يلتزم بها الباحث حتى يتجنب التحيز الشخصي والوقوع في بعض الاخطاء . كما يساعد استخدام الطريقة الإحصائية الى وصول الباحث الى النتائج الدقيقة بأقصر طريق وأقل كلفة .

6-1. المراحل الرئيسية للطريقة الاحصائية

1- تحديد مشكلة البحث

2 - جمع البيانات والمعلومات

3 - تصنيف البيانات وتبويبها

4 - عرض البيانات

5 - حساب المؤشرات أو المعالم للبيانات

6 - التفسير والتنبؤ

7-1. تصنيف وتبويب البيانات ، تكوين الجداول التكرارية البسيطة والمزدوجة

1-7-1. طبيعة البيانات الاحصائية

عند جمع البيانات حول ظاهرة ما نرمز للظاهرة بالرمز **y** او **x** او اي رمز اخر وكل مفردة او مشاهدة ترمز لها $\mathbf{y_i}$ او $\mathbf{x_i}$ فمثلا عند دراسة اطوال الطلبة لأحدى الجامعات فأننا نرمز لصفة الطول (الظاهرة) y ولطول اي طالب (المفردة) بالرمز y_i . هذا وان قيمة $\mathbf{y_i}$ قد تختلف من طالب الى اخر لهذا نقول ان y متغير variables وعليه فان المتغير : هو اي ظاهرة تظهر اختلافات بين مفرداتها ويرمز له بالرمز y او اي رمز اخر . وتنقسم المتغيرات الى قسمين :-

1 - **متغيرات وصفية او نوعية** :- هي الظواهر او الصفات التي لا يمكن قياسها مباشرة بالأرقام العددية مثل صفة لون العين (ازرق ،اسود ،بني) او الحالة الاجتماعية (غني ،متوسط الحال ، فقير) او الجنس (ذكر ،انثى)......الخ .

2 - **متغيرات كمية** : - هي الظواهر او الصفات التي يمكن قياسها مباشرة بأرقام عددية مثل صفة الطول ، الوزن ، العمر ، كمية المحصول .

وتنقسم المتغيرات الكمية الى قسمين هما : -

أ - **متغيرات متصلة (مستمرة)** : المتغير المستمر هو المتغير الذي تأخذ المشاهدة او المفردة فيه اية قيمة رقمية في مدى معين ، فلو فرضنا ان اطوال طلبة جامعة ما تتراوح بين 130,5 سم و170 سم اي ان المتغير y يمكن

ان ياخذ اي قيمة بين 130,5 و170 سم ومثال ذلك الوزن وكمية المحصول ودرجات الحرارة ، وبصورة عامة كل البيانات التي تقاس تعتبر بيانات لمتغير مستمر (تأخذ القيم عدد صحيح او كسر) .

ب ـ **متغيرات غير مستمرة (منفصلة)** : ـ تأخذ المشاهدة او المفردة فيها قيما متباعدة او منقطعة مثال ذلك عدد الثمار او عدد الوحدات الانتاجية او عدد الطلبة في الصفوف الاولى لجامعة ما ، فهي في الغالب تكون اعداد صحيحة . وبصورة عامة كل البيانات التي نحصل عليها من العد تعتبر بيانات لمتغير منفصل .

1-8. المجتمع والعينة :

المجتمع : ـ عبارة عن جميع القيم او المفردات التي يمكن ان يأخذها المتغير ، فلو كانت دراستنا حول اطوال طلبة جامعة ما فان المجتمع في هذه الحالة هو اطوال جميع الطلبة في تلك الجامعة .والمجتمع اما ان يكون مجتمعا محدودا اي ممكن حصر عدد مفرداته . او يكون مجتمعا غير محدودا وهو المجتمع الذي من الصعب حصر عدد مفرداته مثل مجتمع نوع سمك معين في نهر دجلة او عدد البكتريا في حقل ما .

اما **العينة** : فهي جزء من المجتمع . فالعينة عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع ففي بعض الاحيان دراسة المجتمع ككل قد يكون صعبا او يحتاج الى وقت وجهد ومال فيستعاض عن دراسة المجتمع بدراسة العينة وصفاتها ومنها يستنتج خواص المجتمع الاصلي الذي اخذت منه العينة .

1-9. اسلوب تصميم البحوث

هناك اعتبارات كثيرة يتوقف عليها تصميم البحث وعلى الباحث ان يأخذ بنظر الاعتبار مسالة الحصول على البيانات والمعلومات بأقصر وقت واقل جهد واوطئ كلفة وعليه يجب مراعاة مايلي عند تصميم البحث

1 ـ **تحديد الغرض من البحث** : من الضروري ان يكون الهدف محددا بشكل واضح ودقيق معروفة اهدافه واوجه الاستفادة من نتائجه.

2 ـ **امكانية التنفيذ العملي للبحث** : ـ من الضروري تحديد المتطلبات التي تلتزمها عملية تنفيذ البحث كالموارد المالية المطلوبة والامكانيات البشرية المتاحة في تحقيق بعض فقرات البحث وكذلك التأكد من مدى توفر البيانات والمعلومات الدقيقة عن مشكلة البحث.

3 ـ **تحديد اطار البحث** : ـ من المهم ان يحدد الباحث نوع وطبيعة مجال البحث او المجتمع الاحصائي والمجتمع الاحصائي عبارة عن مجموعة وحدات او مفردات ذات صفة او صفات مشتركة فمثلا اذا كان البحث يتعلق بأطوال طلبة جامعة بغداد فان المجتمع الاحصائي هو جميع الطلبة في جامعة بغداد والمفردة الطالب او الطالبة في هذا المجتمع. واذا كان البحث حول دخل العائلة الفلاحية في العراق فالمجتمع الاحصائي هو العوائل الفلاحية

الساكنة في العراق والوحدة الاحصائية او المفردة هي العائلة الواحدة والمجتمع يكون اما مجتمع محدد وهو المجتمع الذي يمكن الوصول الى كل مفردة فيه مثل مجتمع جامعة بغداد او يكون مجتمع غير محدد مثل كريات الدم البيضاء في دم الانسان ومجتمع الاسماك في نهر دجلة.

1-10. اسلوب جمع البيانات والمعلومات : ـ للوصول الى البيانات والمعلومات هناك اسلوبان يمكن من خلالهما جمع هذه البيانات والمعلومات كل منهما له ميزاته وعيوبه وهذان الاسلوب هما : ـ

1 ـ **اسلوب التسجيل الشامل** : ـ هو جمع البيانات من جميع المفردات التي يتكون منها المجتمع مجال البحث ومثال ذلك التعداد العام للسكان من **ميزات** هذا الاسلوب يعطي بيانات كاملة حول الظواهر التي يتم البحث عنها اما **عيوبه** فان هذا الاسلوب يحتاج الى وقت وجهد ومال كما لا يمكن استخدام هذا الاسلوب في المجتمعات غير المحددة.

2 ـ **اسلوب العينات** : ـ هو اخذ وحدات من المجتمع الاحصائي تسمى العينة sample والغرض من اخذ العينة ان تكون بديلا عن المجتمع الاحصائي وعن طريق صفاتها يتمكن الباحث ان يصف خواص المجتمع بتعميم النتائج التي حصل عليها من دراسة العينة . تفضل هذه الطريقة عن طريقة التسجيل الشامل للأسباب الاتية :

1 ـ توفر المال والجهد والوقت اللازم لإجراء البحث.

2 ـ صعوبة اجراء التسجيل الشامل بسبب طبيعة المجتمع فقد يكون المجتمع غير محدد او كبير جدا. ومن **عيوب** هذا الاسلوب فان محاولة التعرف على خواص المجتمع عن طريق دراسة جزء منه ينطوي عليه التضحية في دقة النتائج التي تستخرجها .

وتنقسم العينات الى قسمين رئيسيين هما :

1-11. العينات العشوائية والعينات الغير عشوائية

1-11-1. ـ **العينات العشوائية** : هي مجموعة المفردات المختارة من مجتمع الدراسة وليس للباحث دخل في اختيارها . وللعينات العشوائية انواع عديدة منها

أ ـ العينة العشوائية البسيطة Simple random sample

هي اختيار عينة عشوائية من مجتمع الدراسة بطريقة تعطي المفردات نفس الفرصة في الظهور .ويشترط هنا ان يكون المجتمع متجانس(مشترك في الصفات)فمثلا دراسة اسباب التدخين لدى الاناث نلاحظ ان المجتمع متجانس حيث ان كافة مفردات هذا المجتمع هم من الاناث والصفة المشتركة هي التدخين.

مثال (1)

أراد مدير شركة اختيار لجنة مكونة من 5 أشخاص من بين مجموعة موظفين عددهم 50 موظف، كيف يتم الاختيار بطريقة العينة العشوائية البسيطة؟

الحل:

يقوم المدير بكتابة أسماء جميع الموظفين على بطاقات كل اسم على بطاقة ثم يعمل على خلط البطاقات ووضعها في صندوق ثم يسحب 5 بطاقات بحيث يسحب بطاقة في كل مرة.

ب- العينة المنتظمة (الأسلوبية) Systematic sample

وهي العينة التي يتم اختيارها من مجتمع يكون موزعا على أساس معين، كأن يكون تصاعدياً أو تنازلياً.
ومن أمثلتها اختيار عدد من الصكوك المدفوعة من دفتر الصكوك المتسلسل أو اختيار عدد من المنازل المرقمة في محافظة ما.
وتتم عملية الاختيار بتحديد الزيادة المنتظمة (k) ثم تحديد مفردة البداية التي تكون عادة أقل من (k)، ومن ثم يتم إضافة هذه الزيادة المنتظمة بشكل متسلسل.
وتعطى الزيادة المنتظمة (k) بالقانون

$$\mathbf{k = N/n}$$

حيث **N** : حجم المجتمع　　　　　**n** : حجم العينة

مثال 2 :

إذا كان حجم مجتمع ما هو 1000 مفردة ويراد اختيار عينة حجمها 100 مفردة ، أوجد الزيادة المنتظمة k.

الحل:

$$\mathbf{k = N/n}$$

$$= 1000 / 100 = 10$$

مثال 3 :

يراد اختيار عينة حجمها 100 من مجتمع عدد أفراده 1260 أوجد مقدار الزيادة المنتظمة.

الحل:

$$\mathbf{k = N/n}$$

$$\mathbf{= 1260 / 100 \quad = 12.6}$$

ويتم التقريب للأسفل فيكون الجواب هو 13.

مثال4:

يراد اختيار عينة حجمها 200 مفردة من مجتمع حجمه 4000 مفردة ، كيف يتم ذلك بطريقة العينة المنتظمة؟

الحل:

1) نجد مقدار الزيادة المنتظمة k :

$$\mathbf{k = N/n}$$

$$\mathbf{= 4000 / 200 = 20}$$

2) نختار مفردة البداية وتكون أقل من 20 ولتكن 8.

3) نضيف k بشكل متسلسل

$$8 , 8 +20 , 8 + 20 + 20 , 8 + 20 + 20 + 20 + ...$$

$$8 , 28 , 48 , 68 , ... , 388.$$

مثال 5:

دائرة المحاسبة في شركة تود اجراء دراسة لعملاء الشركة من حيث طرق السداد وحجم الطلبات وكيفية الشحن، إذا تم اختيار العينات بطريقة العينة المنتظمة وكان عدد العملاء المسجلين 12500؛ وتم اختيار حجم العينة ليكون 350 أوجد العينة المطلوبة.

الحل:

$$\mathbf{k = N/n}$$

$$\mathbf{= 12500 / 350}$$

$$= 35.7$$

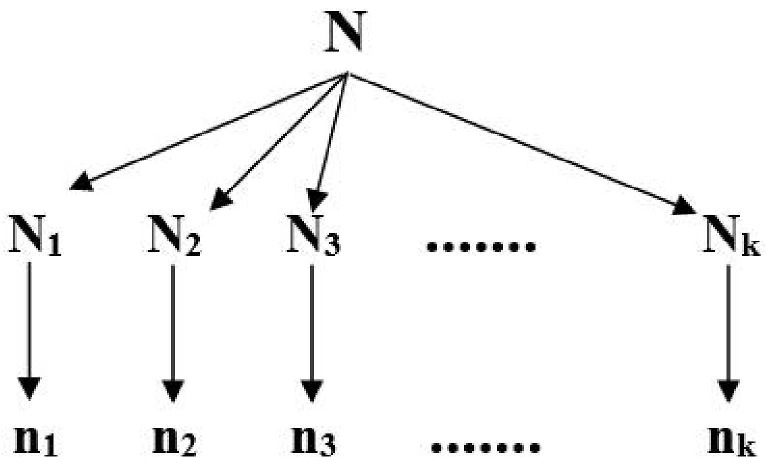
$$\approx 35$$

فتكون عينة العملاء المختارة هي كالآتي؛ إذا اعتبرنا أن مفردة البداية هي 18 :

18 , 53 , 88 , 123 , ...

ج ـ العينة الطبقية العشوائية : ـ يتم اختيار العينة عندما يكون المجتمع غير متجانس ،يقسم المجتمع الى طبقات كل طبقة تعتبر مجتمع متجانس ومن كل مجتمع يتم اختيار عينة عشوائية بسيطة يتناسب حجمها مع حجم الطبقة ثم تجمع هذه العينات ونحصل على الطبقة العشوائية.

مثلا لو كنا بصدد دراسة للمستوى العلمي لإحدى طلبة المعهد التقني نينوى هذا المجتمع غير متجانس من حيث التخصص العلمي فهناك اختصاص ادارة قانونية واختصاص محاسبة واختصاص تقنيات مالية ومصرفية واختصاص سياحة وهكذا.



حيث أن **N** : حجم المجتمع الطبقي

N_i : حجم الطبقة **(i)**.

n_i : حجم العينة من الطبقة (i).

n : حجم العينة المطلوبة.

أي أن

$$N = N_1 + N_2 + N_3 + \dots + N_k$$

$$n = n_1 + n_2 + n_3 + \dots + n_k$$

ونجد حجم العينة من الطبقة (i) بالصيغة

$$n_i = \frac{N_i}{N} \times n$$

مثال 6 :

مجتمع حجمه 10000 مفردة ومكون من 4 طبقات حجم كل طبقة على التوالي :

1000 , 3500 , 4000 , 1500 . يراد سحب عينة حجمها 400 مفردة من هذا المجتمع، كيف يتم ذلك بحيث تمثل هذه العينة تمثيلاً سليماً.

الحل:

$n = 400 \qquad N = 10000$

$N_4 = 1500 \quad N_3 = 4000 \quad N_2 = 3500 \quad N_1 = 1000$

$$n_1 = \frac{N_1}{N} \times n = \frac{1000}{10000} \times 400 = 40$$

$$n_2 = \frac{N_2}{N} \times n = \frac{3500}{10000} \times 400 = 140$$

$$n_3 = \frac{N_3}{N} \times n = \frac{4000}{10000} \times 400 = 160$$

$$n_4 = \frac{N_4}{N} \times n = \frac{1500}{10000} \times 400 = 60$$

نلاحظ أن $n = 40 + 140 + 160 + 60$

$= 400$

وهو حجم العينة المطلوبة.

مثال 7 :

أرادت مجلة متخصصة بحث مقدار الأرباح التي تحققها المنشآت التجارية كلٌ حسب نشاطها التجاري. فعملت بحث الدراسة على 1000 منشأة. فإذا تم اختيار عينة مكونة من 100 منشأة. كيف يتم الاختيار بطريقة سليمة؟

الجدول التالي يبين عدد ونوع كل منشأة في مجتمع الدراسة.

Group	No. Of companies
Financial	300
Diversified	200
Commercial banking	200
Retailing	100
Transportation	100
Utilities	100
Total	**1000**

الحل:

$n = 100 \qquad N = 10000$

$N_6 = 100 \quad N_5 = 100 \quad N_4 = 100 \quad N_3 = 200 \quad N_2 = 200 \quad N_1 = 300$

$$n_1 = \frac{N_1}{N} \times n = \frac{300}{1000} \times 100 = 30$$

$$n_2 = \frac{N_2}{N} \times n = \frac{200}{1000} \times 100 = 20$$

$$n_3 = \frac{N_3}{N} \times n = \frac{200}{1000} \times 100 = 20$$

$$n_4 = \frac{N_4}{N} \times n = \frac{100}{1000} \times 100 = 10$$

$$n_5 = \frac{N_5}{N} \times n = \frac{100}{1000} \times 100 = 10$$

$$n_6 = \frac{N_6}{N} \times n = \frac{100}{1000} \times 100 = 10$$

نلاحظ أن

$$n = 30 + 20 + 20 + 10 + 10 + 10 = 100$$

15-1. الجداول الاحصائية

هناك نوعان رئيسيان من الجداول الاحصائية

1- **الجدول البسيط** : هو الجدول الذي توزع فيه البيانات حسب صفة واحدة ويتألف عادة من عمودين . الاول يمثل تقسيمات صفة الظاهرة الى فئات او مجموعات والثاني يبين عدد المفردات التابعة كل فئة او مجموعة

مثال : الجدول التالي يمثل عدد من الطلبة حسب اوزانهم

فئات الوزن	عدد الطلبة
60-62	5
63-65	15
66-68	45
69-71	27
72-74	8
المجموع	100

2- **الجدول المركب(المزدوج)** : هو الجدول الذي توزع فيه البيانات حسب صفتين او ظاهرتين او اكثر في نفس الوقت ويتألف من :

الصفوف : تمثل فئات او مجاميع احدى الصفتين .

الاعمدة : تمثل فئات او مجاميع الصفة الاخرى.

مثال: الجدول التالي يبين توزيع عدد من الطلبة حسب صفتي الطول والوزن

ا لوزن / **الطول**	51-60	61-70	71-80	**المجموع**
121-140	20	6	4	30
141-160	2	40	10	52
161-180	2	6	10	18
المجموع	24	52	24	100

مثال (1)

لو اردنا عمل توزيع تكراري للأعداد الاتية التي تمثل الوزن بالكيلو غرامات لعشرين طالبا في المعهد التقني نينوى (67,55,65.70,75,60,89,83,65,56,49,65,49,48,69,62,72,45,56,74,)

نتبع الخطوات التالية

1- نبحث عن اكبر قيمة واصغر قيمة في المجموعة وهما (89-45) وذلك للتوصل الى المدى الكلي
2- نجد عدد الفئات
3- نجد طول الفئة
4- نكتب حدود الفئات ونستخرج عدد التكرارات لكل فئة

1- المدى

المدى الكلي = اكبر قيمة – اقل قيمة +1

$$T.R = XL - XS + 1 = 89 - 45 + 1 = 45$$

2- عدد الفئات

$$m = 1 + 3.322 \log n = 1 + 3.322 \log 20 = 5.32 \approx 5$$

او نستخدم القانون التالي:

$$5\,m = 2.5\ \sqrt[4]{n} = 2.5 \times \sqrt[4]{20} = 2.5 \times 2.114 = 5.28 \approx 5$$

3- طول الفئة

$$L = \frac{T.R}{m} = \frac{45}{9} = 9$$

كتابة حدود الفئات واستخراج عدد التكرارات لكل فئة

class	f_i	f_i
40-49	||||	4
50-59	|||	3
60-69	|||||||	7
70-79	||||	4
80-89	||	2
Total	20	20

ملاحظة : هناك عدة طرق لكتابة حدود الفئات :

1- اما ان تكون الاعداد لمتغيرات منفصلة كما في المثال السابق
2- او تكون الاعداد لمتغيرات متصلة وهو الذي يمثل بعدد صحيح او كسر مثل الاوزان والاطوال وتكتب الفئات كالاتي :

من 40 الى اقل من 50

من 50 الى اقل من 60

من 60 الى اقل من 70

وللاختصار تكتب بالصيغة الاتية :

40-

50-

60-

70 -

تمتاز هذه الطريقة بالوضوح وتستخدم غالبا لإعداد التي تمثل متغيرات متصلة

3- وقد تكتب الفئات حسب الصيغة التالية :

اكبر من 40 واقل من 50

اكبر من 50 واقل من 60

اكبر من 60 واقل من 70

وللاختصار تكتب

40-

50-

60-

70-

قد يكون التوزيع في الجدول التكراري البسيط توزيعا منتظما كما في المثال السابق وذلك لتساوي طول الفئة ، او يكون التوزيع غير منتظم اذا كان طول الفئة غير متساوي ،او يكون الجدول مغلقا اذا كان الحد الادنى والحد الاعلى للفئة معروف ، او يكون الجدول مفتوحا في الحالات الاتية :

أ- يكون مفتوحا من الطرف الادنى فقط
ب- يكون مفتوحا من الطرف الاعلى فقط
ت- يكون مفتوحا من الطرفين (اذا كان الحد الانى والحد الاعلى للفئة غير معلوم)

1-16. الجداول التكرارية المزدوجة

مثال (2) : البيانات الاتية لظاهرتين x و y المطلوب تفريغها في جدول تكراري مزدوج

X 2 , 10 , 11 ,4 ,20 , 15 ,15 , 3

Y 3 , 2 ,5 , 6 ,8 ,10 ,2 ,10

X 25 ,25 , 20 ,22 ,15 ,20 ,30 ,30 ,

Y 2 ,6 ,5 ,9 ,15 ,12 ,3 ,9

X 35 , 30 ,35 ,31

Y 10 ,12 ,11 ,4

الحل : نستخرج معلومات لكل ظاهرة على حدى

1- المتغير X

المدى :

$$T.R = XL - XS + 1 = 35 - 2 + 1 = 34$$

عدد الفئات :

$$m = 2.5\sqrt[4]{n} = 2.5 \times \sqrt[4]{20} = 2.5 \times 2.114 = 5.28 \approx 5$$

طول الفئة :

$$L = \frac{T.R}{m} = \frac{34}{5} = 6.8 \approx 7$$

2- **المتغير** Y

3- **المدى** :

$$T.R = XL - XS + 1 = 15 - 2 + 1 = 14$$

4- **عدد الفئات** :

$$m = 1 + 3.322 \log n = 1 + 3.322 \log 20 = 5.31 \approx 5$$

5- طول الفئة :

$$L = \frac{T.R}{m} = \frac{14}{5} = 2.8 \approx 3$$

f_X	f_Y	f_{iX}		f_{iY}	
2-8	2-4	|||	3	||||||	6
9-15	5-7	|||||	5	||||	4
16-22	8-10	||||	4	||||||	6
23-29	11-13	||	2	|||	3
30-36	14-16	||||||	6	|	1
Total		**20**		**20**	

بعد ذلك نضع المعلومات في جدول مزدوج وكالاتي :

X \ Y	2-8	9-15	16-22	23-29	30-36	Total
2-4	|	||		|	||	6
5-7	|	|	|	|		4
8-10	|	|	||		||	6
11-13			|		||	3
14-16		|				1
Total	**3**	**5**	**4**	**2**	**6**	**20**

1-17. التوزيع التكراري المتجمع

التوزيع التكراري البسيط يعطينا عن عدد المفردات في كل فئة لكن في بعض الاحيان نرغب في معرفة عدد المفردات التي قيمتها أقل او أكثر من قيمة معينة في التوزيع التكراري . ويعرف التوزيع التكراري المتجمع: بانه التوزيع الذي يبين كمية التكرار المتجمع عند قيمة معينة من قيم المتغير العشوائي. وهناك نوعان من الجداول التكرارية المتجمعة . ويرمز له ب Fi

أ- **التوزيع التكراري المتجمع الصاعد** : وهو عبارة عن تجميع التكرارات من الفئة الاولى وانتهاء بالفئة الاخيرة منه ويتم حساب التكرارات المتجمعة على اساس الحدود العليا للفئات .

مثال (3) : لو اردنا عمل جدول التوزيع التكراري المتجمع الصاعد للأعداد الاتية التي تمثل الوزن بالكيلو غرامات لعشرين طالبا في المعهد التقني نينوى

(67,55,65.70,75,60,89,83,65,56,49,65,49,48,69,62,72,45,56,74,)

فالحل يكون كالاتي:

نجمع التكرارات من الفئة الدنيا الى الفئة العليا

الفئات	التكرارات f_i	الحدود العليا للفئات	المتجمع الصاعد F
40-49	4	اقل من 49	4
50-59	3	اقل من 59	7
60-69	7	اقل من 69	14
70-79	4	اقل من 79	18
80-89	2	اقل من 89	20
المجموع	**20**		

ب- **جدول التوزيع المتجمع النازل** : عبارة عن تجميع التكرارات ابتداء من الفئات العليا وانتهاء بالفئات الدنيا ،بعبارة اخرى تناقص التكرارات ابتداء بالفئة الاولى في التوزيع وانتهاء بالفئة الاخيرة منه . ويتم حساب التكرارات المتجمعة النازلة على اساس الحدود الدنيا للفئات .

فلو رجعنا الى المثال رقم (3) فان التوزيع التكراري المتجمع النازل يكون كالاتي :

الفئات	التكرارات fi	الحدود الدنيا للفئات	المتجمع النازل F
40-49	4	40 فاكثر	20
50-59	3	50 فاكثر	16
60-69	7	60 فاكثر	13
70-79	4	70 فاكثر	6
80-89	2	80 فاكثر	2
المجموع	**20**		

الأسئلة

س1/ ما هو الهدف من دراسة مادة الإحصاء ؟

س2/ عرف علم الإحصاء وما هي اقسامه ؟

س3/ ما هي أهمية علم الإحصاء وما هي علاقته بالعلوم الأخرى ؟

س4/ ما المقصود بالمتغيرات المستمرة وغير المستمرة ؟

س5/ ما المقصود بالعينة والمجتمع ؟

س6/ ما هو أسلوب تصميم البحث ؟

مقاييس النزعة المركزية
Central Tendency Measurement

بعد أن استعرضنا في محاضراتنا السابقة أساليب جمع وتصنيف وتبويب البيانات وكيفية تمثيلها في جداول ورسوم هندسية وبيانية الهدف منها إعطاء صورة واضحة وسريعة تبين ماهية هذه البيانات . أما ألان فسوف نتطرق الى كيفية تمثيل مجموعة من البيانات بقيمة واحدة فقط من خلال مقياس يدعى مقياس النزعة مركزية أو مقياس متوسط هذه القيمة وهذا العدد يميل لان يقع في وسط تلك المجموعة من البيانات في حال ترتيبها حسب صغرها أو كبرها . أي أن هذا العدد يؤول لان يتمركز في وسط المجموعة التي احتسب منها . وأهم مقاييس التوسط هي :

1 – الوسط الحسابي The Arithmetic Mean
2 – الوسيط The Median Mean
3 – المنوال The Mode
4 – الوسط الهندسي The Geometric Mean
5 – الوسط التوافقي The Harmonic Mean

هذه المقاييس جميعا تبحث عن قيمة تتمركز حولها أغلبية هذه البيانات برقم واحد يعبر عن أو يمثل جميع بيانات تلك المجموعة .وسنشرح كيفية حساب كل مقياس من المقاييس أعلاه في حالتين :
أ – بيانات مبوبة .
ب – بيانات غير مبوبة .

1.2 – الوسط الحسابي *The Arithmetic Mean*

وهو أهم مقاييس النزعة المركزية لما يمتاز به من خصائص وسهولة حسابه كما انه من أكثر مقاييس النزعة المركزية استخداما وهو عبارة عن مجموع القيم مقسوما على عددها وعادة ما يرمز له ($\bar{X}, \bar{Y}, \bar{Z} \ldots$)
طرق حسابه :-

- حساب الوسط الحسابي للبيانات غير المبوبة :

إذا كان المتغير X_i يأخذ القيم $X_1, X_2, X_3, \ldots X_n$ فان الوسط الحسابي لهذه القيم يرمز له بالرمز $\bar{X}$ ويعبر عنه على انه :

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

مثال (1.2)
أوجد الوسط الحسابي لأجور 10 عمال كانت أجورهم الشهرية (بآلاف الدنانير) كما يلي :
الحل :
600 ، 750 ، 450 ، 380 ، 360 ، 520 ، 440 ، 525 ، 425 ، 550

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{5000}{10} = 500$$ ألف دينار

- حساب الوسط الحسابي للبيانات المبوبة :

يمكن حساب الوسط الحسابي للبيانات المبوبة باستخدام القانون التالي :

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i X_i}{\sum_{i=1}^{n} f_i}$$

حيث أن f_i تمثل التكرارات وان X_i تمثل مراكز الفئات .
ويتم الحل على النحو التالي :
1 – نحسب مراكز الفئات .
2 - نحسب $f_i X_i$ لكل فئة (حاصل ضرب مراكز الفئات بالتكرارات المقابلة لها) .
3 – نطبق القانون .

مثال (2.2):

من جدول التوزيع التكراري التالي اوجد الوسط الحسابي :

الفئات	التكرارات f_i	مراكز الفئات	x f_i
03 – 07	8	5	40
08 – 12	3	10	30
13 – 17	2	15	30
18 – 22	3	20	60
23 - 27	4	25	100
	20		260

الحل

الفئات	التكرارات f_i	مراكز الفئات X_i	$f_i\, x_i$
03 – 07	8	5	40
08 – 12	3	10	30
13 – 17	2	15	30
18 – 22	3	20	60
23 - 27	4	25	100
	20		**260**

$$\bar{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{260}{20} = 13$$

1.1.2 - الوسط الحسابي المرجح :

بعض الظواهر تختلف قيمتها من حيث الأهمية النسبية تسمى هذه الأهمية النسبية وزن أو ترجيح لكل مفردة وعند حساب الوسط الحسابي وعند حساب الوسط الحسابي لتلك القيم يسمى بالوسط الحسابي المرجح أو الموزون والذي يرمز له بالرمز $\bar{X}_W$ حيث W تمثل الأوزان الترجيحية وتكون صيغته كما يلي :

$$\bar{X}_W = \frac{\sum_{i=1}^{n} X_i W_i}{\sum_{i=1}^{n} W_i}$$

مثال (3.2) :

الجدول التالي يبين درجات أحد طلبة قسم العلوم المالية والمصرفية في 9 مواد دراسية وعدد الوحدات الأسبوعية لتلك المواد والمطلوب حساب معدل الطالب في تلك المواد مرجحا بعدد الوحدات :

اسم المادة	عدد الوحدات W_i	الدرجة X_i
مبادئ إحصاء	3	70
مبادئ إدارة	3	65
مبادئ اقتصاد	3	50
مبادئ محاسبة	3	70
ثقافة مصرفية	3	82
نصوص E	2	63
حاسوب	2	72
حقوق إنسان	2	77
لغة عربية	2	79
	23	$\sum$

الحل :

X_iW_i	عدد الوحدات W_i	الدرجة X_i
210	3	70
195	3	65
150	3	50
210	3	70
246	3	82
126	2	63
144	2	72
154	2	77
158	2	79
1593	**23**	$\sum$

$$\bar{X}_W = \frac{\sum_{i=1}^{n} X_i W_i}{\sum_{i=1}^{n} W_i} = \frac{1593}{23} = 69.2$$

2 .1.2 - مزايا وعيوب الوسط الحسابي :

- **مزايا الوسط الحسابي :**

1. هو نقطة اتزان المشاهدات ويمتاز بسهولة حسابه وبساطة فكرته .
2. خضوعه للعمليات الحسابية .
3. إن حسابه يستند الى كافة البيانات المتاحة .

- **عيوب الوسط الحسابي**

1. يتأثر بالقيم المتطرفة والقيم الشاذة لذا لا يصلح للتوزيعات التكرارية الملتوية .
2. لا يصلح في حالة الفئات المفتوحة (لعدم وجود مركز فئة) .
3. لا يمكن حسابه في المتغيرات النوعية إلا إذا أمكن تحويل التغير النوعي الى كمي .
4. لا يمكن حسابه في حالة فقدان قيمة أو أكثر لعدم إمكانية حساب المجموع الكلي للقيم لأنه يعتمد على جميع القيم في حسابه .

خواص الوسط الحسابي :[1]

أ – مجموع انحرافات القيم عن وسطها الحسابي تساوي صفر أي :

للبيانات غير المبوبة:

$$\sum (X_i - \bar{X}) = 0$$

ويمكن برهنته كما يلي :

$$\sum (X_i - \bar{X}) = \sum X_i - n\bar{X}$$

$$= \sum X_i - n\left(\frac{\sum X_i}{n}\right)$$

$$= \sum X_i - \sum X_i = 0$$

ولتوضيح ذلك في البيانات غير المبوبة نأخذ المثال التالي :

مثال (4.2) :

$X_i = 8$ ، 12 ، 6 ، 9 ، 5

$$\bar{X} = \frac{40}{5} = 8$$

فلو تم طرح الوسط الحسابي من كل قيمة من القيم الأصلية :

X_i	$X_i - \bar{X}$
8	0
12	4
6	− 2
9	1
5	− 3
40	0

للبيانات المبوبة :

2 – انظر الى :

- د.خاشع الراوي ، مبادئ الإحصاء ، مطابع جامعة الموصل ، ص ص 68 – 74 .
- د. محمود حسن المشهداني وأمين حنا هرمز ، الإحصاء ، بغداد ، ص ص162 – 166.
- د. ثائر فيصل شاهر ،الإحصاء في العلوم الإدارية والمالية ، دار الحامد للنشر والتوزيع ، عمان 2010 ، ص ص 84 -92 .

$$\sum f_i (X_i - \bar{X}) = 0$$

$$\sum f_i (X_i - \bar{X}) = \sum f_i X_i - \bar{X} \sum f_i$$

$$= \sum f_i X_i - \left(\frac{\sum f_i X_i}{\sum f_i} \right) \sum f_i$$

$$= \sum f_i X_i - \sum f_i X_i = 0$$

مثال (5.2) :

للتحقق من تلك الخاصية في البيانات المبوبة نأخذ البيانات التالية التي وسطها الحسابي 13:

الفئات	التكرارات f_i	مراكز الفئات X_i	$X_i - \bar{X}$	$f_i\left(X_i - \bar{X}\right)$
03 –07	8	05	-8	-64
08 – 12	3	10	-3	- 9
13 – 17	2	15	2	4
18 – 22	3	20	7	21
23 - 27	4	25	12	48
	20			0

ب – عند إضافة عدد ثابت (k) الى كل قيمة من قيم المشاهدات فأن الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الأصلية + العدد الثابت (k) :

$$Y_i = X_i + k$$

$$\bar{Y} = \bar{X} + k$$

البرهان :

$$Y_i = X_i + k$$

$$\sum Y_i = \sum (X_i + \bar{X})$$

$$\sum Y_i = \sum X_i + nk$$

$$\frac{\sum Y_i}{n} = \frac{\sum X_i}{n} + \frac{nk}{n}$$

$$\bar{Y} = \bar{X} + k$$

مثال (6.2) : باستخدام بيانات المثال (4.2) والتي وسطها الحسابي = 8 وللتحقق من تلك الخاصية نضيف العدد 3 الى كل مفردات العينة وكما يلي :

X_i	$X_i + 3$
8 12 6 9 5	11 15 9 12 8
$\sum X_i = 40$ $\bar{X} = 8$	$\sum X_i = 55$

$$\bar{X} = \frac{55}{5} = 11$$

والذي يعني الوسط الحسابي القديم + 3 (أي 8 + 3 = 11)

ج – عند طرح عدد ثابت (k) من كل قيمة من قيم المشاهدات فأن الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الأصلية - العدد الثابت (k) :

$$Y_i = X_i - k$$

$$\bar{Y} = \bar{X} - k$$

البرهان :

تطبق نفس الخطوات السابقة مع تغيير الإشارة من موجب الى سالب .

مثال (7.2) :

باستخدام بيانات المثال (4.2) والتي وسطها الحسابي = 8 وبطرح العدد 2 من كل مفردات العينة :

X_i	$X_i - 2$
8 12 6 9 5	6 10 4 7 3
$\sum X_i = 40$ $\bar{X} = 8$	$\sum X_i = 30$

$$\bar{X} = \frac{30}{5} = 6$$

أي أن $\quad \bar{X} - 2 = 8 - 2 = 6$

د – إذا ضربت كل قيمة من قيم المشاهدات في قيمة ثابتة (k) فأن الوسط الحسابي للقيم الجديدة يساوي = الوسط الحسابي للقيم الأصلية × العدد الثابت k أي أن :

$$Y_i = kX_i$$

$$\bar{Y} = k\,\bar{X}$$

البرهان :

$$Y_i = kX_i$$

$$\sum Y_i = k \sum X_i$$

$$\frac{\sum Y_i}{n} = \frac{k \sum X_i}{n}$$

$$\bar{Y} = k\,\bar{X}$$

مثال (8.2) : باستخدام بيانات المثال (4.2) والتي وسطها الحسابي = 6 وبضرب العينة

بالعدد 4 :

X_i	X_i
8 12 6 9 5	32 48 24 36 20
$\sum X_i = 40$ $\bar{X} = 8$	$\sum X_i = 160$

$$\bar{X} = \frac{160}{5} = 32$$

ويلاحظ أن الـوسط الحسابي الجديد = 32 هو عبارة عن 8 ×4 وهذا ما يحقق تلك الخاصية .

هـ ـ الوسط الحسابي لمجموع قيم متغيرين = مجموع الوسطين الحسابيين للمتغيرين :

$$Z_i = X_i + Y_i$$

$$\bar{Z} = \bar{X} + \bar{Y}$$

البرهان :

$$Z_i = X_i + Y_i$$

$$\sum Z_i = \sum (X_i + Y_i)$$

$$\sum Z_i = \sum X_i + \sum Y_i$$

$$\frac{\sum Z_i}{n} = \frac{\sum X_i}{n} + \frac{\sum Y_i}{n}$$

$$\bar{Z} = \bar{X} + \bar{Y}$$

مثال (9.2) :

باستخدام نفس العينة في المثال (2.4) والتي وسطها الحسابي = 8 وعينة أخرى كما في الجدول أدناه ووسطها الحسابي = 7 يمكن برهنة أعلاه رقميا كما يلي :

X_i	Y_i	$Z_i = X_i + Y_i$
8	7	15
12	11	23
6	5	11
9	8	17
5	4	9
$\sum X_i = 40$ $\bar{X} = 8$	$\sum X_i = 35$ $\bar{X} = 7$	$\sum Z_i = 75$

$$\bar{Z} = \frac{\sum Z_i}{n} = \frac{75}{5} = 13$$

الوسيط: *The Median*

أحد مقاييس النزعة المركزية وهو القيمة التي تتوسط مجموعة القيم بعد ترتيبها تصاعديا أو تنازليا ويرمز له بالرمز (Me) .

أ – حساب الوسيط للبيانات غير المبوبة :

يمكن حساب الوسيط للبيانات غير المبوبة بالخطوات الآتية :

- يتم ترتيب القيم تصاعديا أو تنازليا .
- إذا كان عدد القيم فرديا فأن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$.
- أذا كان عدد القيم زوجيا فأن الوسيط عبارة عن الوسط الحسابي للقيمتين الواقعتين في الترتيبين $\frac{n}{2}$ و $\frac{n}{2}+1$.

مثال (10.2) :

الأتي أوزان 11 فرادا والمطلوب حساب متوسط وزنهم باستخدام الوسيط .

73 ، 69 ، 82 ، 66 ، 85 ، 67 ، 59 ، 77 ، 80 ، 65 ، 71

- نرتب القيم تصاعديا :

59 ، 65 ، 66 ، 67 ، 69 ، 71 ، 73 ، 77 ، 80 ، 82 ، 85

- نجد ترتيب الوسيط AMe : $AMe = \frac{11+1}{2} = 6$

$$Me = 71$$

مثال (11.2) :

ألأتي درجات 8 طلاب استخدم الوسيط لتقدير وزن الطالب .

76 ، 68 ، 90 ، 45 ، 57 ، 78 ، 89 ، 44

الحل :

نرتب القيم تصاعديا :

44 ، 45 ، 57 ، 68 ، 76 ، 78 ، 89 ، 90

ترتيب الوسيط AMe : $\frac{n}{2} = \frac{8}{2} = 4$ $\quad \frac{n}{2}+1 = \frac{8}{2}+1 = 5$

$$Me = \frac{68+76}{2} = 72$$

ب – حساب الوسيط للبيانات المبوبة :

يعتمد حساب الوسيط للبيانات المبوبة على التوزيع التكراري المتجمع الصاعد أو النازل ويتم استخراجه بتطبيق القانون التالي :

$$Me = L_1 + \left[\frac{\sum f_i \big/ 2 - f^*}{f_m} \right] W$$

حيث أن :

L_1 = الحد الأدنى الحقيقي لفئة الوسيط .

$\sum f_i$ = مجموع التكرارات .

f^* = التكرار المتجمع الصاعد السابق لفئة الوسيط (عند بداية فئة الوسيط) .

f_m = تكرار فئة الوسيط (أي التكرار المتجمع الصاعد عند نهاية فئة الوسيط - التكرار المتجمع الصاعد السابق لفئة الوسيط).

W: طول الفئة

ويمكن تلخيص خطوات الحل كما يلي :-

- تكوين جدول توزيع تكراري تجميعي تصاعدي أو تنازلي .
- إيجاد ترتيب الوسيط (موقع الوسيط) AMe باستخدام القانون : $AMe = \frac{\sum f_i}{2}$.
- البحث عن القيمة $\frac{\sum f_i}{2}$ (ما يساويها أو اكبر منها) .
- نحدد فئة الوسيط وهي الفئة التي تقع تلك القيمة بين حديها وذلك عن طريق إيجاد قيمتين متتاليتين في التكرار التجميعي التصاعدي يقع بينهما ترتيب الوسيط ، يقابل هاتين القيمتين حدا فئة الوسيط الأدنى والأعلى (ويستحسن أخذ الحدود الحقيقية لهذه الفئة) ، فإذا كانت قيمة الوسيط مساوية لأي تكرار متجمع صاعد فأن فئة ذلك التكرار ستكون هي الفئة الوسيطية ، أما إذا وقعت بين تكرارين متجمعين فإن الفئة اللاحقة لقيمة ترتيب الوسيط ستكون هي فئة الوسيط [1]أي الفئة التي تقابل التكرار المتجمع الصاعد الأكبر من بينهما [2].

---

[1] - د عبد الحميد عبد المجيد البلداوي ، أساليب الإحصاء ، مصدر سابق ، ص 75 .

[2] - د.ثائر فيصل شاهر ، مصدر سابق ،ص 97 .

مثال (12.2) :
أوجد الوسيط للتوزيع التكراري التالي :

الفئات	التكرارات f_i
03 – 07	8
08 – 12	3
13 – 17	2
18 – 22	3
23 - 27	4
	20

الحل :
- نستخرج الحدود الحقيقية للفئات والتكرار التجميعي التصاعدي:

الفئات	التكرارات f_i	الحدود الحقيقية للفئات	التكرارالتجميعي التصاعدي
03 – 07	8	02.5 – 07.5	8
08 – 12	3	07.5 – 12.5	11
13 – 17	2	12.5 – 17.5	13
18 – 22	3	17.5 – 22.5	16
23 - 27	4	22.5 – 27.5	20
	20		

نستخرج ترتيب الوسيط : $AMe = \frac{\sum f_i}{2} = \frac{20}{2} = 10$

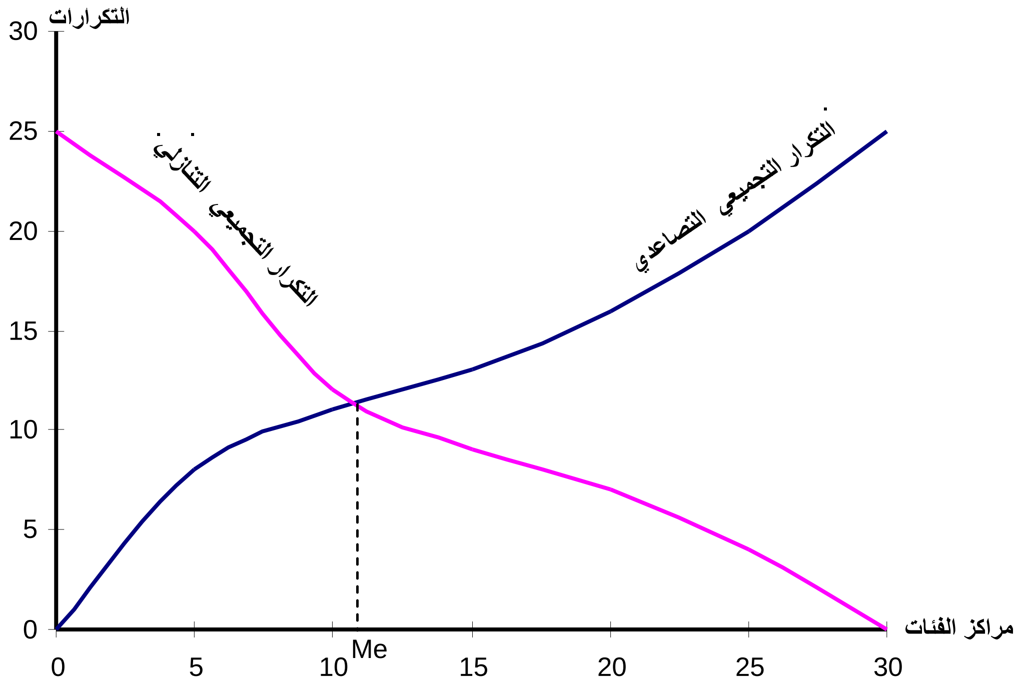
- إن الوسيط يقع بين التكرارين التصاعديين 8 و 11 .
- f_m تكرار فئة الوسيط (3= 8- 11) .
- $f^* =$ التكرار المتجمع الصاعد السابق لفئة الوسيط (عند بداية فئة الوسيط) = 8
- L_1 الحد الأدنى الحقيقي لفئة الوسيط $L_1 = 07.5$.
- نحسب قيمة الوسيط وفق الصيغة التالية :

$$Me = L_1 + \left[\frac{\sum f_i / 2 - f^*}{f_m} \right] W$$

$$Me = 7.5 + \left[\frac{20/2 - 8}{3}\right](5) = 10.8$$

وبنفس بيانات المثال السابق يمكن أيضا إيجاد ترتيب قيمة الوسيط هندسيا للمنحنيين التصاعدي والتنازلي وذلك بإنزال عمود من نقطة تقاطعهما الى المحور الأفقي ليقطعه في نقطة هي قيمة الوسيط كما مبين في الشكل البياني التالي :

شكل (2 – 1) استخراج قيمة الوسيط هندسيا



ويلاحظ بأن قيمة الوسيط المستخرجة بيانيا هي ذات القيمة المستخرجة بموجب المعادلة السابقة .

كما يمكن استخراج الوسيط من الجداول التكرارية المفتوحة والمثال التالي يوضح ذلك .

مثال (13.2) :

أوجد الوسيط من جدول التكرار التالي الذي يمثل القروض الممنوحة من احد المصارف التجارية في بغداد خلال عام 2011 لعدد من المزارعين بملايين الدنانير:

عدد المزارعين	مبالغ القروض
50	5 - أقل من 1
33	06 – 10
65	11 – 15
29	16 – 20
31	21 - 25
16	26 - 30 فأكثر
224	

الحل :

بتطبيق نفس خطوات حل المثال السابق نصل على :

$$Me = L_1 + \left[\frac{\sum f_i /_2 - f^*}{f_m} \right] W$$

التكرار التجميعي الصاعد	الحدود الحقيقية للفئات	عدد المزارعين	مبالغ القروض
50	4.5 – أقل من 0.5	50	5 - أقل من 1
83	5.5 – 10.5	33	06 – 10
148	10.5 – 15.5	65	11 – 15
177	15.5 – 20.5	29	16 – 20
208	20.5 – 25.5	31	21 - 25
224	25.5- 29.5 فأكثر	16	26 - 30 فأكثر
		224	

$$Me = 10.5 + \left[\frac{112 - 83}{65} \right] 5 = 12.73$$

2.2. 1- مميزات وعيوب الوسيط :

- مميزات الوسيط :

1. سهولة حسابه سوءا كانت البيانات مبوبة أو غير مبوبة .
2. يمكن حسابه من جداول التوزيع التكراري المفتوحة.

3. يمثل الوسيط البيانات تمثيلا سليما حينما تكون القيم للمفردات المرتبة متقاربة .
4. يمكن تقدير الوسيط في حالة الصفات الوصفية التي لا تقاس بأعداد مباشرة شريطة أن تكون قابلة للترتيب .
5. مجموع الانحرافات المطلقة لأفراد العينة عن قيمة الوسيط أقل من أي قيمة أخرى .
6. يمكن تعينه بيانيا .

- **عيوب الوسيط** :

1. ليس الوسيط شائع الاستعمال كالوسط الحسابي بالرغم من سهولة فهمه فهو اقل استعمالا منه .
2. لا يمكن الاستفادة من الوسيط حسابيا ومثال ذلك انه لا يمكن حساب وسيط عام لعدد من العينات المعروف قيمة وسيط كل منها كما هو الحال في الوسط الحسابي .
3. ليس الوسيط حساسا للتغيرات التي تحدث في قيم المفردات الداخلة في التوزيع ، فقد نتمكن من تغيير هذه القيم دون أن تتأثر قيمة الوسيط طالما أن هذا التغيير لا ينقله الى الجهة الأخرى .
4. لا تستعمل جميع قيم المتغير في حسابه بل يعتمد على جزء منها .

المنوال: *The Mode*

يعرف المنوال لمجموعة من القيم بأنه القيمة الأكثر شيوعاً بينها ، وقد يوجد أكثر من منوال لمجموعة القيم ، كما إن هذا المقياس يعد من المقاييس المستخدمة في الحالة العملية بكثرة فمثل هذه القيمة نجدها في مقاسات الملابس إذ توجد بعض المقاسات تتكرر لأغلبية الناس أو في مقاسات الأحذية .

أ ـ حساب المنوال للبيانات غير المبوبة :

إذا كان لدينا n من المشاهدات فان المنوال لهذه المشاهدات هو المشاهدة او القيمة الأكثر تكرارا بين هذه المشاهدات ويرمز له بالرمز Mo .

ويجب ملاحظة ما يلي :

1 ـ في حال وجود أكثر من منوال (متجاورة) فمتوسط هذه المنوالات يعتبر منوال التوزيع حال وجود تلك القيمة في التوزيع.

2 ـ يسمى التوزيع أحادي المنوال إذا وجد منوال واحد، وثنائي المنوال إذا وجد منوالان ، وهكذا.

3 ـ إذا وجد أكثر من منوال فالأكثر تكرار بينهم يدعى المنوال الرئيسي *(Mode Major)* والأخرى تعرف بالمنوال الفرعي أو الثانوي *(Mode Minor)* .

1 ـ قد لا يكون هناك منوال أو لا يوجد منوال للمشاهدات .

مثال (14.2) :

مجموعة القيم : 10 ، 13، 8 ، 9 ، 10 ، 10، 23 لها منوال واحد هو القيمة 10 .

مجموعة القيم : 12 ، 13، 7 ، 7 ، 12 ، 12، 23 يوجد أكثر من منوال 12، 7 .

مجموعة القيم : 12 ، 5 ، 4 ، 18 ،10، 16 ، 8 لا يوجد منوال لها .

ب ـ حساب المنوال للبيانات المبوبة :

يستخرج المنوال للبيانات المبوبة بتطبيق القانون التالي :

$$Mo = L_1 + \left[\frac{d_1}{d_1 + d_2} \right] W$$

حيث أن :

فئة المنوال تلك الفئة التي تمتلك أكبر التكرارات .

وان : L_1 = الحد الأدنى الحقيقي لفئة المنوال .

d_1 = الفرق بين تكرار فئة المنوال والفئة السابقة لها .

d_2 = الفرق بين تكرار فئة المنوال والفئة اللاحقة لها .

W = طول الفئة .

مثال (15.2) :

أوجد المنوال لجدول التوزيع التكراري التالي :

الفئات	التكرارات f_i
06 – 10	7
11 – 15	10
16 – 20	18
21 – 25	9
26 – 30	13
	57

الحل :

الفئات	التكرارات f_i
05.5 – 10.5	7
10.5 – 15.5	10
15.5 – 20.5	18
20.5 – 25.5	9
25.5 – 30.5	13
	57

- فئة المنوال هي 20.5 – 15.5 لأنها تقابل اكبر التكرارات وعليه فأن $L_1 = 15.5$

- $8 = 10 - 18 = d_1$

- $9 = 9 - 18 = d_2$

- $5 = W$

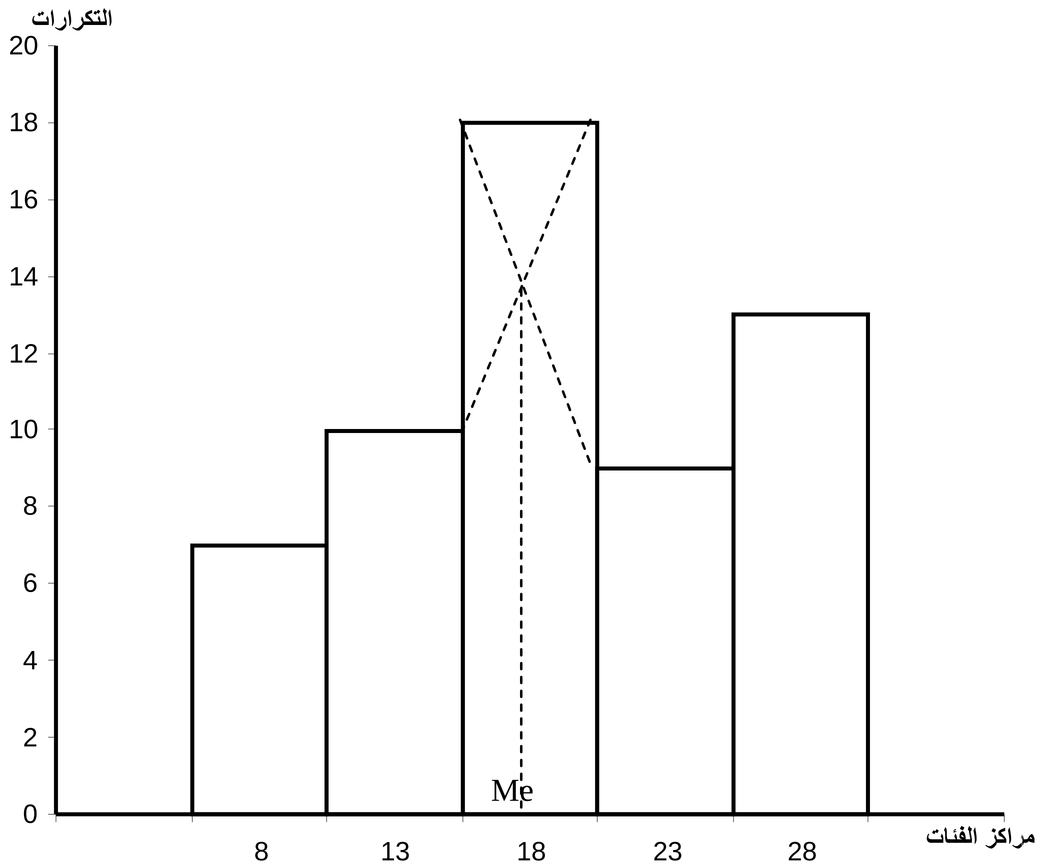
ثم:

$$Mo = L_1 + \left[\frac{d_1}{d_1 + d_2} \right] W$$

$$Mo = 15.5 + \left[\frac{8}{8+9} \right] 5 = 17.85$$

ويمكن كذلك تقدير قيمة المنوال هندسيا ، فمن خلال بيانات المثال السابق يتم رسم المدرج التكراري واستعمال مستطيل الفئة المنوالية باعتباره يمثل أكبر التكرارات والمستطيلان المجاوران له وكما يلي :

شكل (2 -2) استخراج قيمة المنوال بيانيا



ويمكن استخراج المنوال من بيانات الجداول التكرارية المفتوحة والمثال التالي يوضح ذلك :

مثال (16.2) :

من بيانات المثال (13.2) اوجد قيمة المنوال :

الحل :

- نستخرج الحدود الحقيقية للفئات :

الحدود الحقيقية للفئات	عدد المزارعين	مبالغ القروض
4.5 – أقل من 0.5	50	5 - أقل من 1
5.5 – 10.5	33	06 – 10
10.5 – 15.5	65	11 – 15
15.5 – 20.5	29	16 – 20
20.5 – 25.5	31	21 - 25
25.5- 29.5 فأكثر	16	26 - 30 فأكثر

	224	

- ثم نحتسب قيمة المنوال بنفس الخطوات السابقة :

$$Mo = L_1 + \left[\frac{d_1}{d_1 + d_2} \right] W$$

$$Mo = 10.5 + \left[\frac{32}{32 + 36} \right] 5 = 12.8$$

1 .3.2 - مميزات وعيوب المنوال :

- مميزات المنوال :

1. في حسابه لا تستخدم جميع المفردات في التوزيع .
2. المنوال ليس مقياسا يمكن ضربه في عدد المفردات في المجموعة لينتج المجموع الكلي.
3. لا يتأثر المنوال بالقيم الشاذة أو المتطرفة كبيرة كانت أم صغيرة .
4. يمكن حساب المنوال في حالة الجداول المفتوحة .
5. يسهل تقدير المنوال بمجرد النظر إذا كان عدد المفردات قليلا .
6. يمكن تعينه بيانيا .

- عيوب المنوال :

1. قيمة المنوال تقريبية غير معتمد عليها وتتأثر قيمته بشدة من عينة الى أخرى من نفس المجتمع .
2. لا تتغير قيمة المنوال بأحداث تغيرات في القيم الأخرى ما دام تكرار القيمة لم يتغير .
3. أقل المقاييس للمتوسط في دقة حسابه .
4. تتغير قيمة المنوال عند تقديرها في الجدول التكراري إذا تغير عدد الفئات وأطوالها لنفس البيانات .

العلاقة بين المتوسطات الثلاثة (الوسط الحسابي ،الوسيط ،المنوال):

لاحظ كارل بيرسون من نتيجة تجاربه الكثيرة أن العلاقة بين الوسط الحسابي والوسيط والمنوال إذا كان التوزيع متماثل أو قريب من حالة التماثل أي التوزيعات التي لا يكون التوائها شديدا تكون بالشكل التالي :

$$\bar{X} - Me = \frac{\bar{X} - Mo}{3}$$

هذه العلاقة مهمة جدا أثناء التطبيق عندما يتعذر مثلا حساب الوسط الحسابي للتوزيعات التكرارية المفتوحة في حين يمكن حساب الوسيط والمنوال لها لذا يمكن حساب قيمة أي مقياس من المقاييس الثلاث إذا علم المقياسين الآخرين[1] .

مثال (17.2) :
تعذر في أحد التوزيعات القريبة من التماثل الحصول على قيمة الوسط الحسابي في حين أمكن الحصول على قيمتي Me , Mo حيث كانت قيمة Me = 52 وقيمة Mo = 53 .
جد قيمة الوسط الحسابي .

الحل :

$$\bar{X} - Me = \frac{\bar{X} - Mo}{3}$$

$$\bar{X} = 51.5$$

[1]- د.محمود المشهداني وأمين حنا هرمز ، مصدر سابق ، ص 225 .

الوسط الهندسي: *The Geometric Mean*

الوسط الهندسي عبارة عن الوسط الحسابي للوغاريتمات القيم (وليس القيم نفسها) كما انه يعد من مقاييس النزعة المركزية المهمة لتعدد استعمالاته فهو يستخدم مع النسب ومعدلات النمو ومع الأرقام القياسية ومعدلات الفائدة كونه اقل تأثرا من الوسط الحسابي بتباين حجم القيم المتطرفة .

أ - حساب الوسط الهندسي للبيانات غير المبوبة Ungrouped data :

الوسط الهندسي هو الجذر النوني (n) لمجموعة من القيم التي تتضمنها عينة ما :

$$\bar{G} = \sqrt[n]{X_1 . X_2 ..... X_n}$$

حيث أن :

$X_{1,} X_2 ..... X_n$ تمثل مفردات العينة .

مثال (18.2):

اوجد الوسط الهندسي للقيم التالية : 8 ، 4 ، 2 .

الحل :

$$\bar{G} = \sqrt[3]{(2)(4)(8)} \quad = 4$$

يلاحظ بأنه من الممكن أستخرج الوسط الهندسي بالطريقة أعلاه وذلك لقلة المفردات المراد استخراج الوسط الهندسي لها ولكن عندما يكون هناك عدد كبير من المفردات نستعمل اللوغاريتمات بدلا من القيم الأصلية وكما يلي :

$$Log\,\overline{G} = \frac{1}{n}\sum_{i=1}^{\bar{n}} LogX_i$$

ثم نأخذ العدد المقابل الى اللوغاريتم لنحصل على الوسط الهندسي :

$$\overline{G} = \bar{anti} \rightarrow Log_{10}\left(\frac{1}{n}\sum_{i=1}^{n} LogX_I\right)$$

مثال (19.2) :

اوجد الوسط الهندسي للبيانات التالية :

5 ، 7 ، 6 ، 4 ، 12 ، 10 ، 1 ، 3 ، 4 ، 8 .

الحل :

$$Log\,\overline{G} = \frac{1}{n}\sum_{i=1}^{\bar{n}} LogX_i$$

$$Log\overline{G} = \frac{Log8 + Log4 + Log3 + Log1 + Log10 + Log12 + Log4 + Log6 + Log7 + Log5}{10}$$

$$Log\overline{G} = 0.69857$$

ثم نأخذ العدد المقابل الى اللوغاريتم :

$$\overline{G} = 5$$

وهناك طريقة أخرى للحل من خلال أخذ حاصل ضرب جميع أعداد المفردات في المثال مع بعضها البعض ثم نستخرج اللوغاريتم لها بعدها نقسم الناتج على عدد المفردات أو المشاهدات والبالغة 10 ثم نأخذ العدد المقابل للوغاريتم للحصول على الوسط الهندسي .

ب – حساب الوسط الهندسي للبيانات المبوبة Grouped data :

لاستخراج الوسط الهندسي للبيانات المبوبة نطبق القانون التالي :

$$Log\overline{G} = \frac{\sum_{i=1}^{n} f_i LogX_i}{\sum f_i}$$

حيث أن :

$\sum f_i$ = مجموع التكرارات .

X_i = مراكز الفئات .

مثال (20.2) :

اوجد الوسط لهندسي من الجدول التكراري التالي :

الفئات	التكرارات f_i
03 – 07	8
08 – 12	3
13 – 17	2
18 – 22	3
23 - 27	4
	20

الحل :

الفئات	التكرارات f_i	مراكز الفئات X_i	$LogX_i$	$f_i LogX_i$
03 – 07	8	5	0.6989	5.5912
08 – 12	3	10	1.0	3
13 – 17	2	15	1.1760	2.352
18 – 22	3	20	1.3010	3.903
23 - 27	4	25	1.3979	5.5916
	20			**20.4378**

$$Log\overline{G} = \frac{\sum_{i=1}^{n} f_i LogX_i}{\sum f_i}$$

$$LogG = \frac{20.4378}{20} = 1.02189$$

$$\bar{G} = 10.51$$

1 .5.2 - مميزات وعيوب الوسط الهندسي:

- مميزات الوسط الهندسي :

1. لا يتأثر بالقيم الكبيرة في التوزيع .
2. الوسط الهندسي أصغر دائما من الوسط الحسابي لأي مجموعة من القيم .
3. بساطة فكرته.
4. إن حسابه يستند الى كافة البيانات المتاحة .
5. لا يتأثر بأخطاء المعاينة .
6. يمكن حسابه من التوزيعات التكرارية ذات الفئات المتساوية الأطوال وغير المتساوية .
7. يخضع للعمليات الجبرية .
8. يعطي قيمة أكثر تمثيلا من الوسط الحسابي عند التعامل مع توزيعات شديدة الالتواء نحو اليمين .

- عيوب الوسط الهندسي :

1. لا يمكن حساب الوسط الهندسي لمجموعة من الأعداد إذا احتوت على القيمة صفر أو قيمة سالبة .
2. لا يمكن حسابه من الجداول المفتوحة .
3. لا يمكن تعينه بيانيا .
4. لا يمكن حسابه في حالة فقدان أحدى مفردات العينة أو أكثر .
5. لا يستخدم في حالة كون البيانات وصفية .

الوسط التوافقي: *The Harmonic Mean*

وهو من المقاييس قليلة الاستخدام في التطبيقات الإحصائية ويعرف الوسط التوافقي لمجموعة من القيم بأنه مقلوب الوسط الحسابي لمقلوبات القيم في المجموعة .

أ – حساب الوسط التوافقي للبيانات غير المبوبة: *Ungrouped data*

$$\bar{H} = \frac{n}{\sum \frac{1}{X_i}}$$

مثال (21.2) :

اوجد الوسط التوافقي للقيم التالية :

1 ، 4 ، 2 ، 6 ، 3

الحل :

$$\bar{H} = \frac{n}{\sum \frac{1}{X_i}}$$

$$\bar{H} = \frac{5}{\frac{1}{3}+\frac{1}{6}+\frac{1}{2}+\frac{1}{4}+\frac{1}{1}} = \frac{5}{\frac{27}{12}} = \frac{60}{27} = 2.222$$

ب - حساب الوسط التوافقي للبيانات المبوبة :

الوسط التوافقي للبيانات المبوبة يحسب كما يلي :

$$\bar{H} = \frac{\sum f_i}{\sum \left(\frac{f_i}{X_i} \right)}$$

حيث إن :

X_i = مراكز الفئات .

f_i = التكرارات .

مثال (22.2) :

اوجد الوسط التوافقي لمعطيات الجدول التالي :

الفئات	التكرارات f_i
03 – 07	8
08 – 12	3
13 – 17	2
18 – 22	3
23 - 27	4
	20

الحل :

الفئات	التكرارات f_i	مراكز الفئات X_i	$\frac{f_i}{X_i}$
03 – 07	8	5	1.6
08 – 12	3	10	0.3
13 – 17	2	15	0.133
18 – 22	3	20	0.15
23 - 27	4	25	0.16
	20		**2.343**

$$\bar{H} = \frac{\sum f_i}{\sum\left(\frac{f_i}{X_i}\right)}$$

$$\bar{H} = \frac{20}{2.343} = 8.53 \approx 8.5$$

1.5.2 - مزايا وعيوب الوسط التوافقي :

- مزايا الوسط التوافقي :

1. حسابه يعتمد على كافة البيانات ويتأثر بأي تغيرات تحدث في المفردات .
2. أقل تأثرا بالقيم الكبيرة مقارنة بالوسط الحسابي .
3. يفضل استخدامه عند حساب متوسط المعدلات وخاصة معدلات الإنتاج .
4. يخضع للعمليات الجبرية .
5. الوسط التوافقي دائما اصغر من الوسط الحسابي وأقل استعمالا منه.

- عيوب الوسط التوافقي :

1. لا يمكن حسابه من الجداول التكرارية المفتوحة .
2. لا يمكن حسابه إذا كانت إحدى القيم تساوي صفر .
3. لا يمكن تعينه بيانيا .
4. لا يمكن استخدامه في حالة كون البيانات وصفية .