الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي
جامعة أكلي محند أولحاج
- البويرة -
كلية العلوم الاقتصادية والتجارية وعلوم التسيير

Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique
Université Akli Mohand Oulhadj - Bouira -
Tasdawit Akli Muḥend Ulḥağ - Tubirett -
Faculté des sciences économiques,
commerciales et des sciences de gestion

مطبوعة موجهة للطلبة بعنوان :

محاضرات في الإحصاء 01

مدعومة بأمثلة محلولة

من إعداد : د/ العمري علي

السنة الجامعية : (2024-2025)

مقدمة

إن دراسة الظواهر الاقتصادية تحتاج إلى خطوات تبدأ من جمع بيانات أو معلومات عن الظاهرة وتنتهي بالتنبؤ بنتائجها في المستقبل، فهذه الخطوات تحتاج إلى تنظيم محكم نظرا لإعتمادها على قواعد رياضية وإحصائية، فمن خلال هذه المطبوعة الموسومة بـ " محاضرات في الإحصاء 01" المقدمة لطلبة السنة الأولى جذع مشترك سنعرض هذه الخطوات من خلال مجموعة محاور، فالبداية سنعرض فيها الجانب النظري لعلم الإحصاء والمتغيرات الاحصائية بالإضافة إلى مراحل البحث في دراسة الظواهر، لننتقل بعدها إلى عرض البيانات الإحصائية والذي يبدأ بتحديد نوع المتغيرة ومن خلاله نعتمد على المنهجية الخاصة بها، فتبويب بيانات المتغيرة الكمية المتقطعة والممثلة بعدد أطفال العائلة أو عدد طلبة قسم ... يختلف عن تبويب بيانات متغيرة كمية مستمرة متمثلة في قامات أو أوزان أشخاص أو رقم أعمال مجموعة مؤسسات...، وكذلك تبويب بيانات نوعية (كيفية)، وبعد التبويب الجدولي نحول هذا الأخير إلى تمثيل بياني.

من أجل تحليل البيانات المبوبة (المجدولة) لابد من مقاييس تستخدم في ذلك، وبما أنه يوجد عدد كبير من هذه المقاييس (النزعة المركزية، التشتت ، الشكل والتمركز) واستخداماتها المختلفة، سنقسمها إلى محاور لنبين أنواعها، طرق حسابها واستخداماتها في تحليل البيانات المجمعة.

كما تتضمن هذه المطبوعة محورين هامين هما محور الأرقام القياسية ومحور الارتباط والانحدار، فالرقم القياسي هو مقياس إحصائي يقيس التغير الذي يطرأ على الظواهر والمتغيرات بسبب تأثير عوامل مختلفة، الأمر الذي يؤدي إلى تغيير قيمها من زمن إلى آخر، ومن مكان لآخر، أما الارتباط والانحدار فالأول يقيس العلاقة الارتباطية ما بين متغيرين سواءا كميين أو نوعيين، والثاني أي الانحدار يحدد شكل العلاقة ما بين متغيرين والتي قد تكون خطية أو غير خطية، طردية أو عكسية مع التركيز في نهاية المحور على أهم شكل مستخدم وهو شكل ومعادلة الانحدار الخطي البسيط.

من خلال ما سبق ذكره يمكن القول أن محتوى هذه المطبوعة كان عصارة سنوات من تدريس مقياس الإحصاء 01 لطلبة السنة الأولى جذع مشترك في كلية العلوم الاقتصادية، التجارية وعلوم التسيير ووفق آخر منهاج للوزارة الوصية كما هو موضح في المرفق اللاحق، وبالتالي نتمنى أن يكون هذ العمل مفيدا لهم لاحتوائه على أمثلة تطبيقية.

دليل المادة التعليمية Syllabus

اسم المادة : إحصاء 1

الميدان :	علوم اقتصادية وعلوم تجارية وعلوم التسيير	**الفرع الشعبة :**	جميع الشعب
التخصص :	جذع مشترك	**المستوى :**	الأولى ليسانس
السداسي :	الأول	**السنة الجامعية :**	-

وصف المادة التعليمية

المكتسبات	يحتاج الطالب فقط إلى معرفة أهم العمليات و القواعد الرياضية التي تم التطرق إليها في مرحلة التعليم المتوسط و الثانوي.
الهدف العام للمادة التعليمية	يهدف هذا المقياس إلى وصف مجموعة من البيانات وتنظيمها وتصنيفها وتلخيصها وعرضها بطريقة واضحة في صورة جداول أو أشكال بيانية وحساب المقاييس الإحصائية المختلفة لوصف متغير ما (أو أكثر) في مجتمع ما.
أهداف التعلم (المهارات المراد الوصول إليها)	بعد دراسة مقياس إحصاء 1، سيكون الطالب قادرا على : - التحكم في المفاهيم الأساسية للإحصاء الوصفي. - تلخيص وتبويب البيانات في شكل جداول وتمثيلها بيانيا. - حساب وتفسير مختلف المقاييس الأساسية: مقاييس النزعة المركزية، والتشتت، والشكل، والتمركز. - حساب الأرقام القياسية والتعرف على معناها وفائدتها واستخدامها. - تحليل وتكميم العلاقة بين متغيرين وقياس قوة واتجاه هذه العلاقة.

محتوى المادة التعليمية

المحور الأول	**المفاهيم الأساسية لعلم الإحصاء:** التعريف بعلم الإحصاء وتاريخ نشأته ، التعريف ببعض المصطلحات الاحصائية ، المتغيرة الاحصائية وأنواعها ، مصادر جمع البيانات الاحصائية وطرق جمعها ، أنواع البحوث الاحصائية وخطوات القيام بها.
المحور الثاني	**عرض البيانات الاحصائية :** يقسم هذا المحور إلى العرض الجدولي للبيانات والذي يحتوي على تعريف العرض الجدولي وأنواع الجداول الاحصائية ، كما يحتوي على عرض البيانات الاحصائية للمتغيرة النوعية بنوعيها والمتغيرة الكمية بنوعيها، بالإضافة إلى عرض طريقة حساب التكرارت النسبية والتكرارات المجمعة الصاعدة والنازلة ، أما القسم الثاني من هذا المحور فهو مخصص للعرض البياني للمتغيرة الكمية المتقطعة والكمية المستمرة والمتغيرة النوعية
المحور الثالث	**مقاييس النزعة المركزية :** المتوسط الحسابي، المتوسط الهندسي، المتوسط التوافقي والمتوسط التربيعي. الوسيط وأشباه الوسيط (المئينات، العشيريات والربيعيات)، المنوال. مقارنة بين مقاييس النزعة المركزية.

المحور الرابع	**مقاييس التشتت :** مقاييس التشتت المطلقة (المدى العام، المدى الربيعي، الانحراف المتوسط، التباين والانحراف المعياري). مقاييس التشتت النسبي (معامل الإختلاف النسبي ومعامل الاختلاف الربيعي)
المحور الخامس	**مقاييس الشكل:** حساب العزوم، مقاييس الإلتواء (بيرسون، فيشر،ومعاملات أخرى)، مقاييس التفرطح (بيرسون، فيشر، ومعاملات أخرى)
المحور السادس	**مقاييس التمركز:** منحنى لورنز-Lorenz Curve - مؤشر جيني - Gini Index
المحور السابع	**الأرقام القياسية :** الأرقام القياسية البسيطة، الأرقام القياسية المجمعة، الأرقام القياسية المرجحة.
المحور الثامن	**الارتباط والانحدار:** توزيعات المتغيرات ثنائية التغير (جداول التوافق والتكرارات المشتركة) ؛ الارتباط بين متغيرين كيفيين (معامل الاقتران فآي (Phi) ، كاي مربع (χ^2) ومعامل التوافق) ؛ الارتباط بين متغيرين كميين (معامل الارتباط)؛ الانحدار (سحابة النقاط ، الانحدار الخطي البسيط).

إمضاء مسؤول ميدان التكوين

مسؤول فريق ميدان التكوين

أ.د. فرومي حميد

إمضاء نائب العميد للدراسات والمسائل المرتبطة بالطلبة

2021/11/02

نائب العميد مكلف بالدراسات والمسائل المرتبطة بالطلبة

فهرس المحتويات

فهرس المحتويات

فهرس المحتويات

فهرس المحتويات

فهرس المحتويات

المحور الأول:

الأطر النظرية لعلم الإحصاء والمتغيرات الإحصائية

I-1. تمهيد:

تقوم الأطر النظرية لعلم الإحصاء على مجموعة من المفاهيم الأساسية التي تحدد كيفية جمع البيانات، تبويبها، تحليلها، واستخلاص النتائج منها. ومن بين أهم هذه المفاهيم نجد المتغيرات الإحصائية التي تُعتبر حجر الأساس في أي دراسة تحليلية، إذ تمثل الخصائص أو السمات التي يمكن قياسها لدى الظواهر أو الأفراد، وتتغير قيمها من حالة إلى أخرى. وتنقسم هذه المتغيرات إلى أنواع مختلفة مثل المتغيرات الكمية والكيفية، والمتغيرات المستمرة والمنفصلة، وهو ما يسمح باختيار الأدوات الإحصائية المناسبة لمعالجتها.

كما تُسهم هذه الأطر في توفير منهجية علمية دقيقة تساعد الباحث على فهم العلاقات بين المتغيرات، واختبار الفرضيات، وبناء النماذج الإحصائية التي تُستخدم لتفسير الظواهر وتوقع سلوكها المستقبلي. وبالتالي، يمثّل هذا المحور مدخلاً أساسياً لكل من يرغب في التعمق في الإحصاء وفهم أسسه النظرية وتطبيقاته العملية.

I-2. تعريف علم الإحصاء:

علم الإحصاء هو علم من العلوم التقنية، وردت له عدة تعاريف من بينها :

- أنه علم جمع البيانات في صورة قياسات رقمية، ثم تصنيفها وتبويبها أو تنظيمها وتلخيصها وعرضها بيانيا أو تحليلها رياضيا ووضع الفرضيات وإختبار النتائج وصولا إلا الاستنتاجات [1] .
- هو مجموعة من الطرق العلمية لجمع وتبويب وعرض وتحليل البيانات العددية، هذا التحليل الذي يساعد على إستخلاص نتائج مفيدة وبناء قرارات منطقية [2] .
- هو علم يتعامل مع جمع وتحليل وتفسير وتقديم البيانات [3].
- Statistics كلمة لاتينية مشتقة من كلمة State والتي تعني الدولة، حيث كان الاعتقاد في البداية بأن علم الإحصاء هو علم الدولة أو علم الملوك لأن من يقوم بجمع البيانات أو الاحصائيات هم المسؤولين عن الدولة أو المملكة دون غيرهم، ومع مرور الزمن أصبح يستخدم من طرف الجميع وفي كل المجالات وأصبح يطلق عليه علم الإحصاء.

وبناءا على التعاريف السابقة يمكن القول أن الاحصاء هو علم يهتم بدراسة الظواهر، حيث يبدأ بجمع البيانات عنها، ثم تصنيفها وتبويبها، ثم عرضها بيانيا وتحليلها من أجل إستخلاص نتائج قد تستخدم في إتخاذ قرارات والتنبؤ بنتائج تلك القرارات مستقبلا.

I-3. الهدف من تدريس الإحصاء:

إن تدريس الإحصاء له عدة أهداف أبرزها:

- تحويل البيانات الأولية (الخامة) إلى بيانات أكثر تنظيما وعرضها بيانيا تعطي لها تحليلا دقيقا.
- إتباع الخطوات المهمة في الإحصاء تجعل الباحث يصل إلى نتائج دقيقة عن الظاهرة المدروسة.
- إنطلاقا من تبويب البيانات وتنظيمها وتحليلها يمكن للباحث إتخاذ قرارات مهمة تؤدي إلى نتائج مثالية مستقبلا.

---

[1] - خالد أحمد فرحان المشهداني و رائد عبد الخالق عبد الله العبيدي ، 2013 ، مبادئ الإحصاء : متضمن التحليل الاحصائي spss ، دار الأيام ، عمان، الأردن، ص17.

[2] - محمد بوهزة، 2011، محاضرات في الإحصاء الوصفي، دار المحمدية العامة، الجزائر، ص 7.

[3] -BARBARA ILLOWSKY and SUSAN DEAN ;2013 ; **Introductory Statistics** ; OpenStax ; Texas ; U.S.A ; p05

- مساعدة الباحث على معرفة وفهم العلاقة بين المتغيرات.

I-4. تطبيقات الإحصاء في الاقتصاد وإدارة الأعمال والدراسات المحاسبية:

يستخدم علم الإحصاء في مجالات عديدة أبرزها الاقتصاد وإدارة الأعمال، فهو يستخدم في الدراسات الاقتصادية التي تهدف عادة إلى التنبؤ والتخطيط سواءا كانت على مستوى مشروع صناعي (المستوى الجزئي) أم على مستوى الاقتصاد القومي (الاقتصاد الكلي)، حتى أصبحت المؤشرات والمقاييس الإحصائية من أهم الوسائل العلمية اللازمة في التحليل الاقتصادي، فدراسة الأسعار والأجور والاستثمار والادخار والاستهلاك والتصدير والاستراد وأي منغير آخر من المتغيرات الاقتصادية أصبحت تعتمد إلى درجة كبيرة على الأسلوب الكمي الاحصائي.

أما في ميدان إدارة الأعمال فيسهم الإحصاء في دراسة إتجاهات المبيعات وبيان الآثار الموسمية والدورية لها بهدف إعداد الخطط المستقبلية للمؤسسات والشركات المدروسة ، بالإضافة إلى تتبع حجم المخزون من المنتجات والمواد الخام والوقود وذلك بهدف المحافظة على حجم معين من المخزون، كما أن الإحصاء يستخدم بشكل كبير في دراسة احتياجات المستهلكين ورغباتهم وأذواقهم وبالتالي تقدير المبيعات المستقبلية التي تساعد في إعداد الموازنات الخاصة بالإنتاج والتكاليف والمخزون من المواد المختلفة.

يستخدم كذلك علم الإحصاء في الدراسات المحاسبية مثل تحليل القوائم المالية والبحث عن المؤشرات ومعايير محاسبية والتنبؤ بالأرباح التجارية أو الصناعية وتحديد القيمة المضافة على صعيد المؤسسة أو القطاع الصناعي بشكل عام.[1]

I-5. تقسيمات علم الإحصاء:

ينقسم علم الإحصاء إلى قسمين هما الإحصاء الوصفي والاحصاء الاستدلالي، فالأول يكمل الثاني، حيث:

- **الإحصاء الوصفي**: هو ذلك النوع من الإحصاء يقوم بتقديم مجموعة البيانات على شكل جداول أو رسومات بيانية وغيرها من وسائل العرض البياني.

وبناءا على التعريف السابق يمكن تعريف الإحصاء الوصفي على أنه العلم الذي يهتم بدراسة الظواهر، حيث يهتم بجمع البيانات وتنظيمها في جداول وعرضها بيانيا، أي أنه يقوم بوصف معطيات الظاهرة جدوليا وبيانيا فقط.

- **الإحصاء الاستدلالي**: هو علم يعتمد على التقدير واختبار الفرضيات، حيث يعمل على استخلاص تعميمات عن خواص المجتمع من واقع خواص عينة مأخوذة منه، فبالرغم من إختلافه عن الإحصاء الوصفي من حيث المنهجية والأدوات المستعملة إلا أنه يعتمد أساسا على البيانات والاستنتاجات الأولى للإحصاء الوصفي.

على ضوء تعرفي الإحصاء الوصفي والاحصاء الاستدلالي يمكن القول أن علم الإحصاء ينقسم إلى مرحلتين هما :[2]

- **مرحلة أولى** : جمع للبيانات، عرض وتنظيم النتائج، تلخيص للنتائج (عرض جدولي)، رسم بياني وهو تعبير عن مرحلة أولى في التحليل بواسطة مقاييس النزعة المركزية، مقاييس التشتت....إلخ وهي إستنتاجات أولية ذات طابع وصفي وهي مرحلة يطلق عليها الإحصاء الوصفي.

---

[1] - عدنان عباس حميدان وآخرون، 2016، مباديء الإحصاء، منشورات جامعة دمشق، سوريا، ص ص : 23 ، 24.

[2] - محمد بوهزة، مرجع سابق، ص 8.

- **مرحلة ثانية** : يتم فيها الاستدلال وتعميم النتائج، ثم إكتشاف القوانين المنظمة للظواهر المدروسة وهي مرحلة يطلق الإحصاء الاستدلالي.

I-6. المصطلحات الإحصائية:

يوجد العديد من المصطلحات المتداولة في علم الإحصاء من بينها: المتغيرة، الاحصائيات، المجتمع.....، ولمعرفة متى وكيف نستخدمها لابد من إعطاء تعريفات لكل منها:

المتغيرة الاحصائية: هي الصفة أو السمة أو الخاصية أو الظاهرة التي يريد الباحث دراستها، وهي" الشيء المشترك بين كل الوحدات (العناصر) الإحصائية "[1] حيث تختلف من عنصر إلى آخر مثل: الطول، الوزن، الاستهلاك، عدد الأطفال عند مجموعة من الأسر، لون العيون ....إلخ

الاحصائيات: هي المعلومات أو البيانات المتعلقة بالظواهر أو المتغيرات الإحصائية والتي " قد تكون مقدمة في شكل جداول أو أشكال إحصائية مختلفة سواءا كان هذا التقدم منشورات في كتيبات خاصة أو في مجلات أو دوريات أو وثائق إدارية "[2] أو قد تكون إلكترونية.

المجتمع الاحصائي: هو كل العناصر أو الأفراد الذين ينصب عليهم الاهتمام في دراسة الظاهرة، أو هو " مجموعة المشاهدات والقياسات الخاصة بمجموعة من الوحدات الإحصائية والتي تخص ظاهرة من الظواهر القابلة للقياس مثل مجتمع من الطلبة، مجتمع من الأسر، مجتمع من المؤسسات....إلخ "[3] .

إن المجتمع الاحصائي يمكن أن يكون محدودا أو غير محدود، فالمجتمع المحدود أو النهائي يكون عدد أفراده معروف مثل عدد طلبة جامعة ما، فالطالب في الجامعة له رقم تسجيل معين إداريا وبالتالي العدد الإجمالي للطلبة يكون مضبوطا ومعروفا، أما المجتمع غير المحدود أو اللانهائي فعدد أفراده غير معروف أو غير محدود مثل عدد مرضى السكري في ولاية ما، لأن العدد لا يمكن ضبطه بفعل أن بعض الأفراد يمكن أنهم مرضى لكن لا يعلمون ذلك، كما ينقسم المجتمع الاحصائي كذلك إلى نوعين:

- **مجتمع الهدف:** هم كل الأفراد أو العناصر الذين تنصب عليهم الدراسة، الذين يتميزون بالصفة أو الخاصية المراد دراستها.
- **مجتمع الدراسة:** هم أفراد مجتمع الهدف الذين تم الحصول منهم على بيانات أو إحصائيات عن المتغيرة الدروسة.

التعداد: هي مصطلح أو كلمة ترتبط عادة بمصطلح السكان لتشكل عبارة " التعداد السكاني"، فهي تعني الحصول على معطيات أو بيانات أو إحصائيات متعلقة بالسكان كعدد الأفراد، عدد الوفيات، عدد الغرف، دخل الأفراد، نوع الأمراض لدى الأفراد،.....إلخ، لكن في الأصل هي ليست مقتصرة على السكان بل بظواهر أخرى كالمناخ، التربة وغيرها.

[1] - عبد الرزاق عزوز، 2010، الكامل في الإحصاء، الجزء الأول، ديوان المطبوعات الجامعية، الجزائر، ص 15.

[2] - محمد راتول، 2006، الإحصاء الوصفي، ديوان المطبوعات الجامعية، الجزائر، ص 3.

[3] - جيلالي جلاطو، 2005، الإحصاء، ديوان المطبوعات الجامعية، الجزائر، ط 05، ص 5.

من تقديم مصطلح التعداد يظهر بأنه عملية حصول على إحصائيات، أي حصر كمي لمعطيات الظواهر، ونظرا للتكاليف الباهظة التي يتطلبها عادة ما تقوم به الدولة أو هيئاتها الرسمية، إذن يمكن تعريف التعداد على أنه عملية الحصول على بيانات كمية عن الظواهر وتقوم به هيئة رسمية.

الوحدة الاحصائية: هي الجزء أو العنصر أو مفردة المجتمع الخاضع للدراسة، " فهي قد تكون شيئا حيويا مثل شخص، طالب، موظف...، وقد تكون شيئا ماديا مثل مؤسسة، سيارة، علبة...، كما قد تكون شيئا معنويا مثل فكرة، مذهب..." [1].

العينة Sample: هي جزء من المجتمع[2] ، أو هي عناصر أو وحدات من المجتمع يتم اختيارهم إما :

- عشوائيا وتسمى بالعينة العشوائية وهي ممثلة للمجتمع خير تمثيل وتعتبر أكثر العينات استخداما، حيث يتم اختيار عناصرها بعدة طرق أبرزها القرعة (طريقة الكيس) أو الدولاب أو الأرقام العشوائية....،
- عمديا وتسمى بالعينة العمدية (غير عشوائية) وهي غير ممثلة للمجتمع بصورة مثالية، مما يخول لها أن تكون أقل وأندر العينات استخداما، حيث يتم اختيار عناصرها وفق شروط محددة ومعروفة مسبقا مثل اختيار ممثلي قسم مع شرط أن يكون من الذكور.

I-7. أنواع البيانات وتصنيف المتغيرات:

يوجد نوعين من البيانات:

- **بيانات إبتدائية**: تسمى كذلك أولية أو خام أو غير مبوبة، وهي بيانات متحصل عليها من وحدات الظاهرة دون تنظيمها.
- **بيانات مبوبة**: هي بيانات أولية أو خام تم تنظيمها في جداول إحصائية، عادة ما يتم الحصول عليها من هيئات رسمية، مجلات، دوريات، مواقع إلكترونية مختصة....إلخ.

أما المتغيرات فتصنف إلى متغيرات كيفية وأخرى كمية، ولكل صنف نوعين، حيث : [3]

- **المتغيرات الكيفية**: هي تلك الخصائص التي لا يمكن قياسها أو غير قابلة للقياس، والتي تنقسم إلى نوعين :

1) **المتغيرة الكيفية (النوعية) القابلة للترتيب** : وهي المتغيرة التي لا يمكن قياسها وفي نفس الوقت يمكن ترتيب وحداتها مثل الرتب العسكرية (جندي، رقيب ....) أو تقديرات النجاح (مقبول، حسن، جيد...) .
2) **المتغيرة الكيفية (النوعية) غير القابلة للترتيب** : وهي المتغيرة التي لا يمكن قياسها وفي نفس الوقت لا يمكن ترتيب وحداتها مثل الجنسية (جزائرية، تونسية،...)، أو الحالة العائلية (أعزب، متزوج...).

- **المتغيرة الكمية**: هي المتغيرات التي يمكن قياسها وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام، وهذ النوع من المتغيرة ينقسم إلى قسمين: متغيرة منقطعة (منفصلة) ومتغيرة مستمرة (متصلة)

[1] - عبد الرزاق عزوز، الجزء الأول ، مرجع سابق، ص 15.

[2] - خالد أحمد فرحان المشهداني و رائد عبد الخالق عبد الله العبيدي ، مرجع سابق، ص 29.

[3] - جيلالي جلاطو، مرجع سابق، ص ص: 6، 7.

1. **المتغيرة المتقطعة (المنفصلة)**: هي تلك المتغيرات التي تأخذ أرقاما صحيحة لا يمكن تجزئتها مثل **عدد** الأطفال في العائلة، **عدد** قطع الغيار المنتجة، **عدد** طلبة الأفواج، **عدد**....إلخ.
2. **المتغيرة المستمرة (المتصلة)**: هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المتناهي لهذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات، مثل أجور عمال مؤسسة، قامات أطفال مدرسة،....إلخ.

I-8. مصادر البيانات Sours Of Data:

يوجد مصدرين للبيانات (الاحصائيات) هما : المصادر المباشرة والمصادر غير المباشرة

- **المصادر المباشرة**: تسمى كذلك بمصادر الميدان، وهي أن يقوم الاحصائي أو الباحث بجمع بيانات الظاهرة (المتغيرة) من وحداتها الأصلية إما عن طريق المقابلة، المراسلة أو الاستبيان، والبيانات التي تجمع بهذه الطريقة أي من المصادر المباشرة تسمى بالبيانات الأولية أو الخام.
- **المصادر غير المباشرة**: تسمى كذلك بالمصادر التاريخية، وهي البيانات التي تم تجميعها وتنظيمها سابقا سواءا من طرف هيئات رسمية أو مختصين لأغراض معينة، هذه البيانات عادة ما تكون محفوظة في سجلات، دراسات، مواقع إلكترونية مختصة.

I-9. طرق جمع البيانات:

يتطلب جمع بيانات عن ظاهرة يراد دراستها جهدا ماديا وبدنيا ويستلزم ذلك بعدا زمنيا، فوفق حجم هذا الجهد المتوفر يتم إما التوجه نحو حصول تام للبيانات ويكون ذلك بالحصول على هذه البيانات من كل عناصر أو وحدات الظاهرة والذي يسمى بالحصر الشامل، أو أن يتم الاتجاه نحو عدد محدود من عناصر أو وحدات الظاهرة، أي يتم الحصول على هذه البيانات من عينة، وهذه العملية تسمى بالمعاينة.

- **جمع البيانات عن طريق الحصر الشامل**: تعتبر عملية الحصر (المسح) الشامل أفضل طريقة لجمع البيانات لأنها تأتي بكل المعلومات التي يعبر عنها مجتمع الدراسة، حيث تكون هذه البيانات شاملة ودقيقة، ومن أبرز الدراسات التي يستخدم فيها المسح الشامل هي المتعلقة بالسكان والتي تسمى " التعداد السكاني"، فهي تمثل جمع البيانات الديمغرافية والاقتصادية والاجتماعية عن كافة أشخاص بلد ما خلال فترة محددة ، وللقيام بذلك يتطلب وسائل مادية وبشرية كبيرة جدا، كما يستغرق وقتا طويلا، وبالتالي عادة من يقوم بذلك الدولة وهيئاتها الرسمية، فالحصر (المسح) الشامل له مزايا عديدة من أبرزها دقة البيانات ، أما أهم عيوبه هو ارتفاع تكاليفه.
- **جمع البيانات عن طريق المعاينة**: هي عملية الحصول على بيانات من أفراد أو وحدات العينة، وهي من أكثر الطرق إستخداما بإعتبارها أقل كلفة ويتم الحصول عليها في أقصر وقت بالمقارنة بالمسح الشامل، فإذا تم إختيار عناصر العينة وفق أسلوب علمي سليم فبياناتها تكون دقيقة، وكلما زاد حجم العينة زادت دقة النتائج وأصبحت أقرب إلى نتائج المسح الشامل، ومن بين أهم أسباب إتباع أسلوب المعاينة :

1. وجود مجتمع دراسة غير محدود، وبالتالي إستحالة إجراء الحصر (المسح) الشامل.
2. توفير الجهد والمال وحتى الوقت اللازم للدراسة.

I–10. مراحل البحوث الإحصائية:

من أجل إجراء بحث أو دراسة إحصائية لابد من إتباع مجموعة من الخطوات أو المراحل تكون وفق التسلسل التالي:

أ– تحديد موضوع الدراسة (الظاهرة) وإطارها العام: إن دراسة أي ظاهرة يتطلب تفكير أولي متعلق بمجموعة من التساؤلات يمكن اختصارها في التساؤل: لماذا اخترت دراسة هذه الظاهرة؟ والذي من خلال الإجابة عنه يتم تحديد الهدف والإطار العام للدراسة، والذي يندرج فيه تحديد المجتمع الاحصائي، كذلك المكان والتكلفة اللازمة لذلك والوقت المخصص لهذه الدراسة، بالإضافة إلى الأسلوب المعتمد ووحدة القياس الملائمة.

ب– جمع البيانات (المعطيات) الإحصائية: بعد تحديد الاطار العام للدراسة والذي من خلاله يتم تحديد المجتمع الاحصائي يكون الاختيار ما بين الحصول على البيانات عن طريق الحصر الشامل أو عن طريق المعاينة، كذلك تحديد مصادرها فيما إذا كانت متوفرة (مصادر غير مباشرة) أو غير متوفرة وتستلزم الذهاب مباشرة إلى مصدرها وهو إما المجتمع (الحصر الشامل) أو الاكتفاء بعينة (المعاينة)، بالإضافة إلى تحديد الأسلوب الملائم للحصول عليها (المقابلة، الاستبيان...)، فمرحلة جمع البيانات هي من أهم مراحل البحث، لأن الحصول على بيانات دقيقة يؤدي إلى دراسة ونتائج موثوقة تساعد في اتخاذ القرارات السليمة.

ت– تنظيم وعرض البيانات: إن مصدر الحصول على البيانات قد يكون مباشرا أو غير مباشر، فالمصدر غير المباشر عادة ما يكون هيئة رسمية أو دراسة سابقة، فتكون هذه البيانات منظمة سواءا جدوليا أو بيانيا، أما إذا كانت من مصادرها المباشرة (مجتمع أو عينة) فتكون في شكل خام، وبالتالي يلجأ صاحب الدراسة إلى تصنيفها وتبويبها عن طريق وضعها في مجموعات متجانسة تشترك في صفة واحدة أو عدة صفات، ثم يقدمها في شكل جداول وأشكال مناسبة يسهل قراءتها.

ث– تحليل البيانات واستقراء النتائج: تعطي البيانات المنظمة المعروضة جدوليا أو بيانيا نظرة أولية عن واقع الظاهرة المدروسة، ولفهم جوانبها الأساسية عادة ما تستخدم أدوات إحصائية مناسبة مثل المتوسط الحسابي، الانحراف المعياري.....إلخ من أجل تحليل هذه البيانات والحصول على نتائج الدراسة واستقراء مدلولها واتخاذ القرارات على أساس تلك النتائج.

ج– المحاكاة والتنبؤ بنتائج الظاهرة في المستقبل: تعتبر المحاكاة من أهم الطرق المستخدمة في التوقعات الخاصة بالظواهر، فهي " عملية تمثيل أو نمذجة أو إنشاء مجموعة من المواقف تمثيلا أو تقليدا لأحداث من واقع الحياة حتى يتيسر عرضها والتعمق فيها لاكتشاف أسرارها والتعرف على نتائجها المحتملة عن قرب" [1]، وبالتالي هذه التجربة يمكن من خلالها توقع نتائج مستقبلية للظاهرة المدروسة.

[1] - الموسى عبد الله عبد العزيز، 2001، استخدام الحاسوب الآلي في التعليم، مكتبة القشري، الرياض، ص 582.

المحور الثاني:

عرض

البيانات الإحصائية

II-1.تمهيد:

تعد عملية جمع البيانات الإحصائية أول خطوة يقوم بها الباحث بعد تحديد ما يريد دراسته (الظاهرة المدروسة)، حيث يصادف مجموعة حقائق (بيانات) غير منظمة يصعب إستعابها وفهمها ، ونظرا لتعدد أنواع هذه البيانات من حيث نوعها سواءا كانت رقمية أو غير رقمية (نوعية) يضطر إلى تقسيمها وتصنيفها إلى مجموعات متجانسة في جداول يسهل قراءتها وتمثيلها بيانيا.

II-2. عرض البيانات الإحصائية:

إن عرض البيانات الإحصائية ينقسم إلى عرض جدولي وعرض بياني.

II-2-1. العرض الجدولي للبيانات الإحصائية:

II-2-1-1. تعريف العرض الجدولي: هو عبارة عن وضع المعلومات الإحصائية في جداول نهائية، يحتوي كل منها على عمودين (سطرين) ، يبين العمود (السطر) الأول قيم الظاهرة أو المتغير المدروس ، وتكون هذه القيم على شكل قيم نقطية أو على شكل مجالات ، أما العمود (السطر) الثاني فيحتوي على تكرارات هذه القيم أو المجالات [1]،كما يمكن أن يحتوي العمود (السطر) الأول على صفات (متغير غير رقمي) وعمودها (سطرها) المقابل على تكرارات هذه الصفات

II-2-1-2. أنواع الجداول الإحصائية:

يوجد نوعين من الجداول الإحصائية هي جدول التوزيع الاحصائي البسيط وجدول التوزيع الاحصائي المزدوج.

II-2-1-2-1. جدول التوزيع التكراري البسيطة:

لقد ذكرنا سابقا أن العرض الجدولي هو عبارة عن سطرين أو عمودين ، فالسطر أو العمود الأول يمثل المتغير (X_i) قد يأخذ قيم نقطية أو مجالات أو صفات، أما السطر أو العمود الثاني يمثل تكرارات المتغير (n_i)، والتمثيل البياني لجدول التوزيع التكراري البسيط يمكن أن يأخذ أحد العرضين التاليين :

الجدول رقم (02-01): الشكلين العمودي والأفقي (الأسطر) لجدول التوزيع التكراري البسيط

المتغير X_i	التكرار n_i
x_1	n_1
x_2	n_2
.	.
.	.
.	.
x_k	n_k
المجموع	$\sum n_i$

المتغير X_i	x_1	x_2	.	.	x_k	المجموع
التكرار n_i	n_1	n_2	.	.	n_k	$\sum n_i$

[1] - جيلالي جلاطو، مرجع سابق، ص 11.

II-2-1-2-2. جدول التوزيع التكراري المزدوج:

يستخدم جدول التوزيع التكراري المزدوج عند دراسة ظاهرتين في آن واحد لنفس المجتمع ، مثل دراسة الطول والعمر لعينة من الأطفال، حيث يتم وضع قيم الظاهرة الأولى على الأسطر ، أما قيم الظاهرة الثانية فيتم وضعها على الأعمدة.

نرمز لقيم الظاهرة الأولى بـ (X_i) حيث : (i=1 ; 2 ; 3 ; … ; k)

ونرمز لقيم الظاهرة الثانية بـ (Y_j) حيث (j=1 ; 2 ; 3 ; … ; z)

الجدول رقم (02–02): الشكل العام لجدول التوزيع التكراري المزدوج

X_i \ Y_j	y_1	y_2	y_3	....	y_z	$n_{(i;\bullet)}$
X_1	$n_{(1;1)}$	$n_{(1;2)}$	$n_{(1;3)}$	....	$n_{(1;z)}$	$n_{(1;\bullet)}$
X_2	$n_{(2;1)}$	$n_{(2;2)}$	$n_{(2;3)}$	....	$n_{(2;z)}$	$n_{(2;\bullet)}$
X_3	$n_{(3;1)}$	$n_{(3;2)}$	$n_{(3;3)}$	....	$n_{(3;z)}$	$n_{(3;\bullet)}$
. .	. .	. .	. .	....	. .	. .
X_k	$n_{(k;1)}$	$n_{(k;2)}$	$n_{(k;3)}$	....	$n_{(k;z)}$	$n_{(k;\bullet)}$
$n_{(\bullet;j)}$	$n_{(\bullet;1)}$	$n_{(\bullet;2)}$	$n_{(\bullet;3)}$	....	$n_{(\bullet;z)}$	$\sum_{i=1}^{K} n_{(i;\bullet)} = \sum_{j=1}^{Z} n_{(\bullet;j)}$

II-2-1-2-3. جداول التوزيع التكرارية المغلقة والمفتوحة:

هذين النوعين من الجداول نجدهما في البيانات المبوبة للمتغيرة الكمية المستمرة، حيث جدول التوزيع التكراري المغلق تكون كل حدود فئاته موجودة (معلومة) ، أما جدول التكراري المفتوح فله ثلاث أصناف ، الصنف الأول تكون فئته الأولى بدون حد أدنى (مثلا أقل من 5) ويسمى بجدول التوزيع التكراري المفتوح من الأدنى، أما الصنف الثاني ففئته الأخيرة بدون حد أعلى (مثلا أكثر من أو يساوي 60) ويسمى بجدول التوزيع التكراري المفتوح من الأعلى ، أما الصنف الثالث ففئتيه الأولى والأخيرة مفتوحتين (مثلا أقل من 10 وأكثر من أو يساوي 80) في نفس الجدول ويسمى بجدول التوزيع التكراري المفتوح من الجانبين، ولعرض أشكال هذه الجداول نأخذ مثال عن بيانات مبوبة لمتغيرة كمية مستمرة.

مثال (02–01): لنعتبر لدينا بيانات مبوبة عن متغيرة كمية مستمرة ذات 4 فئات تبدأ من القيمة 5 وطول كل فئة 10 ليتم من خلالها توضيح الشكل العام لجداول التوزيع التكراري المفتوحة والمغلقة.

الحل : الشكل العام لجداول التوزيع التكراري المغلقة والمفتوحة بناءا هذا المثال تكون كما يلي.

الجدول رقم (02–03): الشكل العام لجدول توزيع تكراري مغلق

الفئات	التكرارات
] 15 5]	n_1
] 25 15]	n_2
] 35 25]	n_3
] 45 35]	n_4
المجموع	$\sum n_i$

الجدول رقم (02–04): الشكل العام لجدول توزيع تكراري مفتوح من الأدنى

الفئات	التكرارات
أقل من 15	n_1
] 25 15]	n_2
] 35 25]	n_3
] 45 35]	n_4
المجموع	$\sum n_i$

الجدول رقم (02–05): الشكل العام لجدول توزيع تكراري مفتوح من الأعلى

الفئات	التكرارات
] 15 5]	n_1
] 25 15]	n_2
] 35 25]	n_3
أكثر من أو يساوي 35	n_4
المجموع	$\sum n_i$

الجدول رقم (02–06): الشكل العام لجدول توزيع تكراري مفتوح من الجانبين

الفئات	التكرارات
أقل من 15	n_1
] 25 15]	n_2
] 35 25]	n_3
أكثر من أو يساوي 35	n_4
المجموع	$\sum n_i$

II-2-1-3. تبويب بيانات جداول التوزيع التكرارية:

نعني بالتبويب " وضع البيانات الإحصائية في جداول وذلك بعد تقسيمها حسب صفاتها المشتركة ، والغاية من هذه العملية هو إختصار البيانات إلى أصغر حيز يمكن أن يستوعبها " [1] ، وتختلف طريقة تبويب بيانات جدول توزيع تكراري حسب نوع المتغير، حيث سنعرض طرق تبويب بيانات جدول تكراري بسيط وتبويب بيانات الجدول التكراري المزدوج تأخذ نفس الخطوات.

II-2-1-3-1. تبويب جدول توزيع تكراري بسيط :

تختلف طريقة تبويب بيانات جدول توزيع تكراري بسيط حسب نوع المتغير لذلك نميز الحالات التالية:

II-2-1-3-1-1. تبويب المتغيرات الكمية:

المتغيرات الكمية هي " المتغيرات التي يمكن قياسها رقميا " [2]وهي نوعين، متغيرة كمية متقطعة (منفصلة) مثل بيانات عن عدد الأطفال لدى مجتمع من الأسر أو عدد الأخطاء المطبعية لمجموعة الكتب، أما النوع الثاني فهو متغيرة كمية مستمرة (متصلة) مثل بيانات عن الطول لدى مجتمع من الأشخاص ، أو بيانات عن الدخل لدى مجتمع من العائلات.

تبويب بيانات المتغيرة الكمية المتقطعة :

المتغيرة المتقطعة (المنفصلة) هي التي بياناتها تأخذ قيما رقمية صحيحة فقط مثل عدد الأفراد لدى مجموعة من العائلات، أو عدد التخصصات في مجموعة من الجامعات، ولتبويب هذه البيانات يتم وضع قيم المتغيرة (X_i) في العمود (السطر) الأول مرتبة ترتيبا تصاعديا ، ويتم وضع تكرارات (n_i) هذه القيم في العمود (السطر) الثاني، ولتوضيح ذلك نأخذ المثال الموالي.

مثال (02-02): البيانات التالية تبين عدد أيام التغيب لدى عمال مصنع خلال شهر

1	4	1	2	2	3	3	1	1	4	4	2	4	5	4	4
3	2	1	2	1	5	5	3	4	3	2	5	5	4	4	1

المطلوب: عرض البيانات في جدول توزيع تكراري ؟

الحل: لتبويب البيانات نبدأ أولا بترتيب هذه القيم (عدد أيام الغيابات) تصاعديا في العمود الخاص بالمتغيرة، ثم نحسب تكرارات كل قيمة ونضعها في الخانة المقابلة لها في العمود الثاني كما هو موضح في الجدول الموالي.

الجدول رقم (02-07): توزيع العمال حسب عدد أيام الغياب

المتغيرة (عدد أيام الغياب) X_i	التكرارات (عدد العمال) n_i
1	7
2	6
3	5
4	9
5	5
$\sum$	32

[1] - عدنان عباس حميدان وآخرون ، مبادئ الإحصاء ، منشورات جامعة دمشق، سوريا ، 2016، ص 60.

[2] - محمد حسين محمد رشيد، الإحصاء: الوصفي والتطبيقي والحيوي، دار صفاء، عمان، الأردن، 2007، ص 27.

تبويب بيانات المتغيرة الكمية المستمرة :

المتغيرة الكمية المستمرة (المتصلة) هي المتغيرة التي قد تأخذ قيما حقيقية (صحيحة أو كسرية أو نسبية ...) مثل الوزن عند مجموعة من الأطفال أو فترة (زمن) العمليات الجراحية في مجموعة من العيادات، ولكي يتم تبويب هذه البيانات التي قد لا تتساوى رغما كثرتها يتم وضعها في مجالات (فئات)، هذه الفئات سواءا من حيث عددها وطولها تحدد وفق مجموعة من الخطوات هي حساب فارق البيانات (الفرق بين القيمة الكبرى والقيمة الصغرى) ، ثم استخراج عدد الفئات وبعدهما يتم حساب طول هذه الفئات، ويتم توضيح ذلك من خلال الخطوات المرتبة كالتالي :

1– تحديد المدى (Range): هو مجال انتشار البيانات، ويعبر عنه بالفرق بين أكبر قيمة في البيانات وأصغر قيمة فيها.

المدى = أكبر قيمة للمتغيرة – أصغر قيمة للمتغيرة	$R = X_{\max} - X_{\min}$

2– تحديد عدد الفئات: يستخدم في تحديد عدد الفئات عدة طرق اشهرها طريقة عالم الرياضيات الألماني هاربرت ستيرجس Herbert Sturges (1882-1958) ، وطريقة عالم الإحصاء الاسكتلندي جورج إيدني يول George Udny Yule (1871– 1951) ، حيث :

طريقة ستيرجس (Sturges): تعطى بالصيغة التالية:

$$K = 1 + 3,322\log(n)$$

حيث : k : عدد الفئات   n : عدد القيم   log (.) : اللوغاريتم العشري

طريقة يول (Yule): تعطى بالصيغة التالية:

$$K = 2,5\sqrt[4]{n} = 2,5(n)^{1/4}$$

حيث : k : عدد الفئات   n : عدد القيم

3– تحديد طول الفئة: يتم وفق العلاقة التالية:

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} \qquad L = \frac{R}{K}$$

4– التحقق من المتراجحة : بعد تحديد طول الفئة لابد من تحقق بأن كل البيانات داخل الفئات ويكون ذلك بتحقق هذه المتراجحة.

طول الفئة × عدد الفئات ≥ المدى

5- تحديد حدود الفئات: يتم في هذه المرحلة تشكيل الفئات، حيث تكون بداية الفئة الأولى تساوي أصغر قيمة في البيانات ونهاية الفئة الأخيرة أكبر من أكبر قيمة في البيانات.

6- تحديد عدد القيم أو المشاهدات (التكرارات): بعد تحديد حدود الفئات يتم وضع كل قيمة داخل الفئة الخاصة بها والتي تسمى بعملية تفريغ البيانات، وبالتالي ينتج لكل فئة مجموعة من القيم التي تسمى بالتكرارات (n_i)، ثم يتم التأكد من أن مجموع التكرارات يساوي العدد الإجمالي للقيم.

7- تحديد مراكز الفئات: بعد تبويب بيانات المتغيرة الكمية المستمرة (المتصلة) نجد أننا نتعامل مع مجال وليس قيمة محددة كما هو الحال عند المتغيرة العشوائية المتقطعة (المنفصلة)، ولتخطي هذه المشكلة لابد من إستخراج قيمة محددة وهي مركز الفئة، التي تعبر عن منتصف الفئة والذي يحسب وفق القاعدة التالية:

مركز الفئة = $\dfrac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{2}$

$$C_i = \frac{L_i + L_{i+1}}{2}$$

حيث : L_i : الحد الأدنى للفئة i و L_{i+1} : الحد الأعلى للفئة i.

مثال (02-03): لتكن البيانات الأولية الموالية التي تمثل النفقات الشهرية لعائلات في إحدى الولايات (الوحدة : 10^3دج)

64	54	69	53	59	63	46	46	52	52	45	42	42	29	36	26
73	35	41	40	24	71	56	58	68	39	68	38	65	62	28	62
50	49	64	48	74	47	70	67	20	72	55	66	27	57	57	45
60	59	75	44	44	34	70	43	67	32	62	66	31	61	61	58

المطلوب : قم بتشكيل العرض الجدولي للبيانات وفق طريقة ستيرجس ؟

الحل: بالرغم من أن البيانات هي قيم صحيحة وتعطي الشك أنها متغيرة متقطعة إلا أن أصل المتغيرة (النفقات الشهرية) هو متغيرة مستمرة وليس متقطعة وبالتالي لابد من إتباع الخطوات السابقة في تبويب هذه البيانات.

أولا_ تحديد المدى العام:

$$R = X_{\max} - X_{\min} = 75 - 20 = 55$$

ثانيا_ تحديد عدد الفئات: طلب منا إستخدام طريقة ستيرجس (Sturges) في تحديد عدد الفئات.

$$K = 1 + 3{,}322\log(n) = 1 + 3{,}322\log(64) = 1 + 3{,}322(1{,}806) = 6{,}99 \cong 7$$

عندما نحصل على عدد حقيقي (ليس صحيح) لابد من تقريب الرقم بمعنى أنه مثلا عندما نجد (k=2.5) لا يمكن القول أنه لدينا فئتين ونصف، وبالتالي عدد الفئات التي لدينا في هذا المثال هو (k=7) أي 7 فئات.

ثالثا_ تحديد طول الفئة:

$$L = \frac{R}{K} = \frac{55}{7} = 7{,}85 \cong 8$$

لكي لا تكون هناك تعقيدات في تشكيل الفئات لابد من تقريب طول الفئة إلى عدد صحيح، حيث هذه التقريبات سيتم التأكد منها إن كانت مقبولة أو غير مقبولة بالتحقق من المتراجحة اللاحقة.

رابعا_ التحقق من المتراجحة:

طول الفئة x عدد الفئات ≤ المدى

ومنه $55 < 56 = 7 \times 8$

خامسا_ تحديد حدود الفئات: نبدأ في تشكيل الفئات من الفئة الأولى التي تبدأ من القيمة الصغرى للبيانات ويضاف لها طول الفئة المحسوب لنصل إلى الحد الأعلى لهذه الفئة، أما الفئة الثانية فتبدأ من نهاية الفئة الأولى وتنتهي بإضافة طول الفئة المحسوب، ونواصل في تحديد الحدود الدنيا والعليا للفئات حتى الفئة الأخيرة، وطريقة ما ذكرت يمكن عرضها كما يلي:

- الفئة الأولى: الحد الأدنى للفئة الأولى هو 20، أما الحد الأعلى هو: 20+8=28. ومنه الفئة الأولى هي: $[20-28[$.
- الفئة الثانية: الحد الأدنى للفئة الثانية هو 28، أما الحد الأعلى هو: 28+8=36. ومنه الفئة الأولى هي: $[28-36[$.
- الفئة الثالثة: الحد الأدنى للفئة الثالثة هو 36، أما الحد الأعلى هو: 36+8=44. ومنه الفئة الأولى هي: $[36-44[$.

ونستمر في تحديد الفئات الباقية بنفس الطريقة.

سادسا_ تبويب البيانات (الجدولة): عملية تبويب البيانات هي تشكيل الجدول المكون من المتغيرة (فئات) وعدد التكرارات(n_i) مع إضافة عمود ثالث مساعد يسمى بعمود التفريغ وهو يستخدم خاصة في تبويب المتغيرة المستمرة، وعند عملية تفريغ البيانات لابد من مراعاة أن يكون لكل قيمة فئة واحدة فقط تنتمي إليها والتأكد كذلك من أن مجموع التكرارات يساوي عدد القيم.

الجدول رقم (02-08): توزيع عدد العائلات حسب النفقات الشهرية

النفقات الشهرية (الفئات) X_i	التفريغ	عدد العائلات (التكرار) n_i	مركز الفئة C_i
$[20-28[$	//	4	24
$[28-36[$	//////	6	32
$[36-44[$	////////	8	40
$[44-52[$	//////////	10	48
$[52-60[$	////////////	12	56
$[60-68[$	//////////////	14	64
$[68-76[$	//////////	10	72
$\sum$	–	64	-

II-2-1-3-1-2. تبويب المتغيرة النوعية (الكيفية):

المتغيرة النوعية (الكيفية) هي " المتغيرة التي لا يمكن قياسها رقميا " [1]، وتعبر عن صفات مثل تقديرات النجاح لمجموعة من الطلبة، رتب عسكرية لأفراد من الشرطة ، ألوان عيون مجتمع من الأطفال أو جنسية مجتمع من طلبة جامعة دولية ...وغيرها، وإنطلاقا من هذه الأمثلة نلاحظ وجود نوعين من هذه المتغيرة وهي متغيرة نوعية (كيفية) قابلة للترتيب، أي أن هذه الصفات لديها درجات ترتيبية (كالرتب وتقديرات النجاح)، وأخرى غير قابلة للترتيب، أي لا تأخذ درجات ترتيبية (كالجنسية و لون العيون)، ولكي يتم تبويب بيانات هذا النوع من المتغيرة لابد من مراعاة ترتيب الصفات ترتيبا تصاعديا (من أقل درجة إلى أكبر درجة) إذا كانت المتغيرة قابلة للترتيب ، أما إذا كانت غير ذلك فكتابة الصفات في عمود المتغيرة لا يخضع لأي شرط ، ثم يتم حساب عدد تكرارات كل صفة ويتم وضع ذلك في عمود التكرارت (n_i) ، ولتوضيح ذلك نأخذ المثال الموالي.

مثال (02-04): البيانات الموالية تمثل رتب عسكرية لعينة من أفراد الجيش الجزائري.

نقيب	لواء	ملازم	لواء	ملازم	عقيد
رائد	رائد	ملازم	رائد	نقيب	رائد
نقيب	رائد	ملازم	ملازم	نقيب	ملازم
ملازم	عقيد	نقيب	نقيب	عقيد	ملازم

المطلوب: تبويب البيانات في جدول توزيع تكراري؟

الحل: الرتب العسكرية هي متغيرة عشوائية كيفية (نوعية) وهي من النوع القابل للترتيب، وبالتالي نقوم بترتيب هذه الصفات ترتيبا تصاعديا في عمود المتغيرة ثم نحسب تكرار كل صفة ونضعها في عمود التكرارات، والتبويب مبين في الجدول الموالي.

الجدول رقم (02-09): توزيع عينة أفراد الجيش الجزائري حسب رتبهم العسكرية

المتغيرة (تقديرات النجاح) (X_i)	التكرارات (عدد أفراد الجيش) (n_i)
ملازم (*Sous lieutenant*)	8
نقيب (*Capitaine*)	6
رائد (*Commandant*)	5
عقيد (*Colonel*)	3
لواء (*Général major*)	2
$\sum$	24

II-2-1-3-2. تبويب جدول توزيع تكراري مزدوج:

إن تبويب بيانات جدول توزيع تكراري مزدوج تشبه عملية تبويب الجدول التكراري البسيط وتتبع نفس الخطوات فقط بوجود متغيرين لنفس المجتمع، ولتوضيح ذلك نأخذ المثال الموالي.

[1] - محمد حسين محمد رشيد، مرجع سابق ، ص 26.

مثال (02–05): تمثل البيانات الموالية عدد الأطفال في مجموعة من الأسر والإنفاق الشهري (10^3 دينار) لهذه الأسر

رقم العائلة	عدد أطفالها	إنفاقها الشهري	رقم العائلة	عدد أطفالها	إنفاقها الشهري	رقم العائلة	عدد أطفالها	إنفاقها الشهري
01	5	48	09	1	21	17	3	34
02	1	26	10	3	35	18	2	30
03	1	24	11	3	36	19	4	40
04	3	36	12	4	42	20	4	41
05	2	32	13	5	48	21	5	46
06	5	47	14	1	20	22	1	22
07	4	42	15	2	31	23	2	28
08	4	40	16	4	39	24	3	33

المطلوب: بوب البيانات بإستخدام طريقة يول، ثم أعطي قراءة لبيانات الجدول ؟

الحل: لدينا نوعين من المتغيرة ، متغيرة كمية متقطعة (عدد أطفال العائلة) ومتغيرة كمية مستمرة (الانفاق الشهري للعائلة) ولكي نقوم بتبويب هذه البيانات نبدأ بخطوات تبويب المتغيرة الكمية المستمرة وبعد ذلك المتغيرة الكمية المتقطعة.

- تبويب بيانات المتغيرة الكمية المستمرة (الانفاق الشهري) :

➢ تحديد المدى العام: $R = X_{\max} - X_{\min} = 48 - 20 = 28$

➢ تحديد عدد الفئات بإستخدام طريقة يول:

$$K = 2{,}5\sqrt[4]{n} = 2{,}5\sqrt[4]{24} = 2{,}5(24)^{\frac{1}{4}} = 2{,}5(24)^{0{,}25} = 5{,}53 \cong 6$$

➢ تحديد طول الفئة:

$$L = \frac{R}{K} = \frac{28}{6} = 4{,}66 \cong 5$$

➢ تحديد من المتراجحة: طول الفئة x عدد الفئات ≤ المدى

ومنه $6 \times 5 = 30 > 28$

إذن تبويب البيانات في جدول التوزيع التكراري المزدوج يكون كما يلي:

الجدول رقم (02–10): تبويب بيانات جدول التوزيع التكراري المزدوج للمتغيرتين (عدد الأطفال والانفاق الشهري لـ24 أسرة)

النفقات / عدد الأطفال	]25 20]	]30 25]	]35 30]	]40 35]	]45 40]	]50 45]	المجموع
1	4	1	0	0	0	0	5
2	0	1	3	0	0	0	4
3	0	0	2	3	0	0	5
4	0	0	0	1	5	0	6
5	0	0	0	0	0	4	4
المجموع	4	2	5	4	5	4	24

قراءة لبيانات الجدول: يظهر من بيانات الجدول أنه كلما زاد عدد أطفال الأسرة إرتفع الانفاق الشهري للعائلة، فمثلا العائلات التي عدد أطفالها واحد لم يتجاوز إنفاقها 30 ألف دينار في حين أن العائلات التي عدد أطفالها اثنان إنفاقها تراوح ما بين 25 و 35 ألف دينار، أما العائلات التي لها أكبر عدد من الأطفال (05) فإنفاقها هو الأعلى.

II–2–1–4. أنواع التكرارات : هناك أنواع عديد للتكرارات وكل منها يتم استخدامه في مجال معين، ومن هذه الأنواع التكرارات المطلقة ، التكرارات التجميعية والتكرارات النسبية.

أ_ التكرارات المطلقة: هي تمثل عدد المفردات أو المشاهدات ويرمز لها بـ (n_i).

ب_ التكرارات المتجمعة (التجميعية): هي التكرارات التراكمية ولها نوعان هما التكرارات التجميعية الصاعدة ($n_i \uparrow$) والتكرارات التجميعية النازلة ($n_i \downarrow$) ، فالتكرار التجميعي الصاعد لأي صفة أو قيمة نقطية أو فئة هو تكرار تلك القيمة أو الصفة أو الفئة مضافا إليه مجموع تكرارات القيم أو الصفات أو الفئات السابقة، أما التكرار التجميعي النازل لأي صفة أو قيمة نقطية أو فئة هو مجموع التكرارات مطروحا منه تكرارات القيم أو الصفات أو الفئات السابقة.

ج _ التكرارات النسبية: يوجد عدة أنواع من التكرارات النسبية منها التكرار النسبي البسيط (f_i)، التكرار النسبي المئوي ($f_i\%$)، التكرار النسبي التجميعي الصاعد ($f_i \uparrow$)، التكرار النسبي التجميعي النازل ($f_i \downarrow$)، التكرار النسبي المئوي الصاعد ($f_i\% \uparrow$) والتكرار النسبي المئوي النازل ($f_i\% \downarrow$)، حيث يتم حساب التكرار النسبي والنسبي المئوي وفق القواعد التالية والتكرارات النسبية والنسبية المئوية التجميعية الصاعدة والنازلة وفق طرق حساب التكرار المطلق التجميعي الصاعد والنازل والذين سنبين طرق حسابهم وفق المثال اللاحق.

➢ التكرار النسبي البسيط (المطلق) :

التكرار النسبي = $\frac{\text{تكرار الفئة أو القيمة أو الفئة}}{\text{إجمالي التكرارات}}$

$$f_i = \frac{n_i}{\sum n_i}$$

التكرار النسبي المئوي = التكرار النسبي x 100

$$f_i\% = f_i \times 100$$

مع العلم أن : $\sum f_i = 1$ و $\sum f_i\% = 100\%$

مثال (02–06): لتكن البيانات الموالية التي تمثل تكرارات متغيرة عشوائية نوعية (تقديرات النجاح لعدد من الطلبة)

المتغيرة (x_i)	التكرارات (n_i)
مقبول	8
حسن	6
جيد	10
جيد جدا	12
ممتاز	4
المجموع	40

المطلوب:

1/ أوجد كل من التكرارات التجميعية الصاعدة والنازلة، التكرارات النسبية والنسبية المئوية والنسبية التجميعية الصاعدة والنازلة ؟

2/ ما هو عدد الطلبة الذين تقديرات نجاحهم على الأكثر جيد جدا ، وما هي نسبتهم ؟

3/ ما هو عدد الطلبة الذين تقديرات نجاحهم على الأقل جدا وما هي نسبتهم ؟

الحل : نستخدم الجدول السابق في حساب التكرارات بكل أنواعها.

الجدول رقم (02–11): طرق حساب أنواع التكرارات

المتغيرة x_i	التكرارات n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i\%$	$f_i \uparrow$	$f_i \downarrow$	$f_i\% \uparrow$	$f_i\% \downarrow$
مقبول	8	8	40	0,2=40÷8	20=100*0.2	0.2	1	20	100
حسن	6	14 = 6+8	32=8-40	0,15=40÷6	15	0.35	0.8	35	80
جيد	10	24=10+14	26=6-32	0,25=40÷10	25	0.6	0.65	60	65
جيد جدا	12	36=12+24	16=10-26	0,3=40÷12	30	0.9	0.4	90	40
ممتاز	4	40=4+36	4=12-16	0,1=40÷4	10	1	0.1	100	10
المجموع	40	/	/	1	100	/	/	/	/

➢ عدد الطلبة الذين تقديرات نجاحهم على الأكثر جيد جدا ونسبتهم يحسب بإحدى الطريقتين:

ط01 : نستخدم التكرارات المطلقة 12 + 10 + 6 + 8 = 36 ، أما نسبتهم فهي 36 ÷ 40 = 0.9 ، أي 90 % من الطلبة.

ط02 : نستخدم التكرارات التجميعية الصاعدة $n_i \uparrow$ المقابلة للصفة جيد جدا وهي 36 ، أما نسبتهم فهي تستخرج من التكرار النسبي المئوي الصاعد $f_i\% \uparrow$ المقابلة للصفة جيد جدا وهي 90 ، أي 90 % من الطلبة.

➢ عدد الطلبة الذين تقديرات نجاحهم على الأقل جدا ونسبتهم يحسب بإحدى الطريقتين:

ط01 : نستخدم التكرارات المطلقة 10 + 12+ 4 = 26 ، أما نسبتهم فهي 26 ÷ 40 = 0.65 ، أي 65 % من الطلبة.

ط02 : نستخدم التكرارات التجميعية النازلة $n_i \downarrow$ المقابلة للصفة جيد وهي 26 ، أما نسبتهم فهي تستخرج من التكرار النسبي المئوي الصاعد $f_i\% \downarrow$ المقابلة للصفة جيد وهي 65 ، أي 65 % من الطلبة.

د ــ التكرارات المعدلة: هي التكرارات التي يتم حسابها عند بيانات مبوبة لمتغيرة كمية مستمرة (فئات) من أجل غرضين هما حساب رسم المدرج التكراري وحساب المنوال وهذا إذا كانت أطوال الفئات غير متساوية، وقاعدة حساب التكرارات المعدلة هي كالتالي:

التكرار المعدل = $\frac{\text{تكرار الفئة}}{\text{طول الفئة}}$ x طول الفئة المختار

$$n_i^* = \frac{n_i}{L_i} \times L^*$$

حيث : طول الفئة المختار (L^*) يكون أصغر طول فئة.

ولتوضيح ذلك نأخذ المثال الموالي.

مثال (02-07): لتكن البيانات المبوبة الموالية والتي تمثل الدخل الشهري لمجموعة من الأفراد، والمطلوب هو تعديل التكرارات

الفئات	التكرارات (n_i)	طول الفئات (L_i)	التكرارات المعدلة (n_i^*)
]30 – 20]	5	10	$n_1^* = \frac{n_1}{L_1} \times L^* = \frac{5}{10} \times 10 = 5$
]50 – 30]	10	20	$n_2^* = \frac{n_2}{L_2} \times L^* = \frac{10}{20} \times 10 = 5$
]60 – 50]	7	10	$n_3^* = \frac{n_3}{L_3} \times L^* = \frac{7}{10} \times 10 = 7$
]80 – 60]	12	20	$n_4^* = \frac{n_4}{L_4} \times L^* = \frac{12}{20} \times 10 = 6$
]100 – 80]	8	20	$n_5^* = \frac{n_5}{L_5} \times L^* = \frac{8}{20} \times 10 = 4$
المجموع	42	/	/

الحل: من أجل تعديل التكرارات نبدأ أولا بتحديد طول الفئة المختار، حيث نحدد طول كل الفئات ونختار الطول الأقل وهو 10، ثم نطبق القاعدة السابقة، فمثلا نحسب التكرار المعدل للفئة الأولى كما يلي.

$$n_1^* = \frac{n_1}{L_1} \times L^* = \frac{5}{10} \times 10 = 5$$

ونواصل حساب التكرارات المعدلة كما هو موضح في الجدول السابق.

II-2-2. العرض البياني للبيانات الإحصائية:

II-2-2-1. تعريف العرض البياني: يقصد بالعرض البياني إستخدام الرسوم للتعبير عن البيانات العددية، وتفيد هذه الطريقة في إظهار البيانات وملاحظة التغيرات فيها بشكل يجذب الانتباه [1] ويساعد في عملية تحليل الظاهرة المدروسة.

II-2-2-2. أنواع العرض البياني:

توجد عدة أنواع للعرض البياني (التمثيلات البيانية) التي تختلف حسب نوع المتغير المدروس

II-2-2-2-1. العروض البيانية في حالة متغير كمي منفصل (متقطع):

للمتغير العشوائي المنفصل ثلاث تمثيلات بيانية هي التمثيل البياني الخاص بالتكرارات البسيطة (المطلقة) والتمثيل البياني للتكرارات التجميعية الصاعدة والتمثيل البياني للتكرارات التجميعية النازلة.

🙰 **العرض البياني للتكرارات المطلقة**: هو أعمدة بسيطة للقيم النقطية للمتغيرة، حيث طول كل عمود يتناسب مع التكرار المقابل للقيمة النقطية ولتوضيح هذا العرض البياني نأخذ المثال الموالي.

مثال (02-08): معطيات الجدول الموالي تمثل عدد الغرف لدى مجموعة من العائلات.

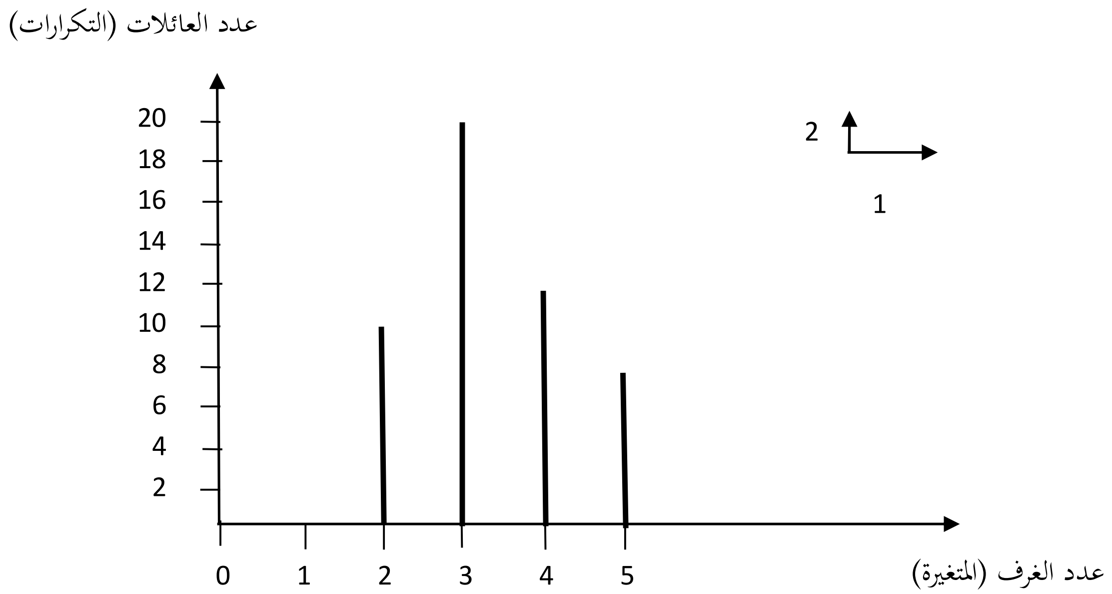
عدد الغرف X_i	2	3	4	5	$\sum$
عدد العائلات n_i	10	20	12	8	50

المطلوب: مثل بيانيا معطيات الجدول ؟

[1] - عدنان عباس حميدان وآخرون، مرجع سابق، ص 64.

الحل: التمثيل البياني لمعطيات الجدول والتي يستخدم فيها قيم المتغيرة مع تكراراتها هي الأعمدة البسيطة.

الشكل رقم (02-01): العرض البياني للتكرارات المطلقة (الأعمدة البسيطة)



- **العرض البياني للتكرارات التجميعية الصاعدة:**

هي قطع مستقيمة أفقية تنطلق من قيمة المتغيرة وتنزل إلى القيمة الأقل منها، أما إرتفاعها فيكون وفق قيمة التكرار المقابل لقيمة المتغيرة.

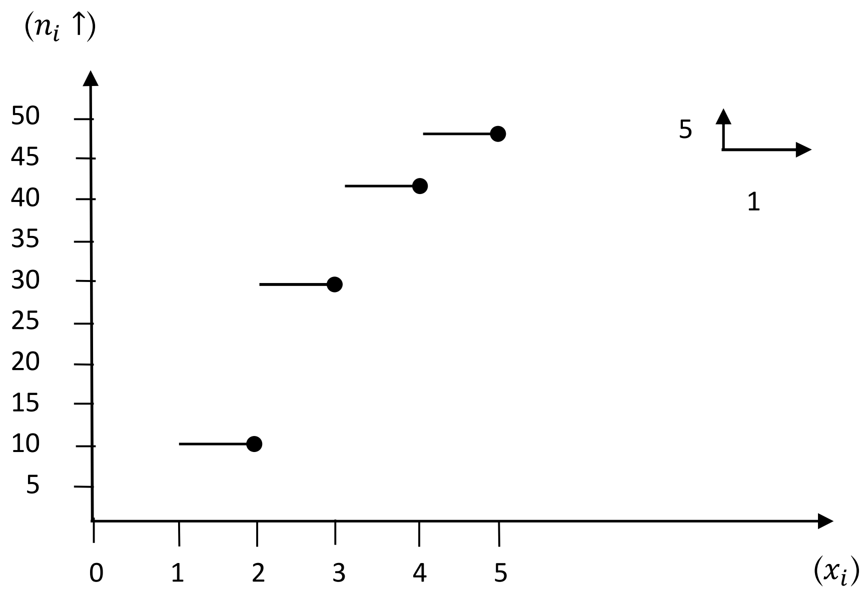
- **العرض البياني للتكرارات التجميعية النازلة:**

هي قطع مستقيمة أفقية تنطلق من قيمة المتغيرة وتصعد إلى القيمة الأكبر منها، أما إرتفاعها فيكون وفق قيمة التكرار المقابل لقيمة المتغيرة.

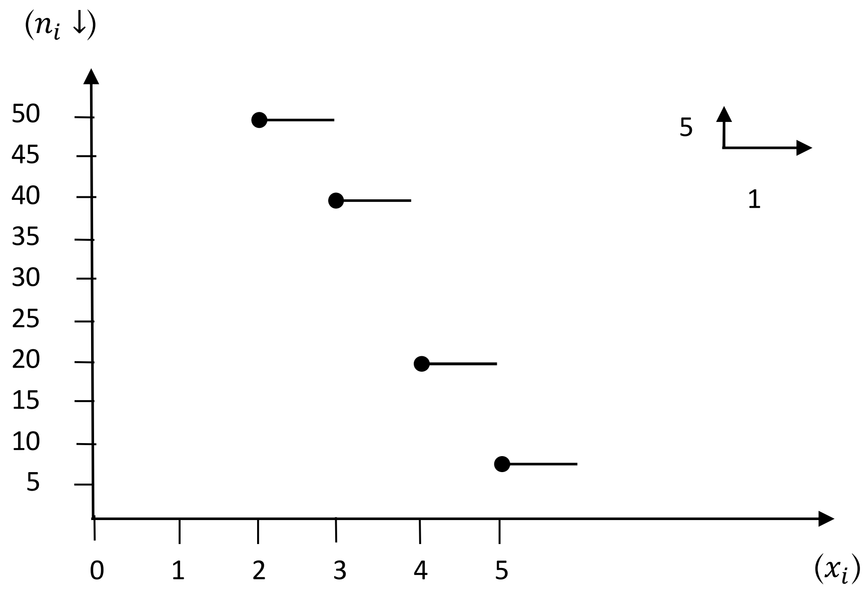
لتوضيح طريقة تمثيل التكرارات التجميعية الصاعدة والنازلة نستخدم معطيات المثال (02-08) السابق.

المتغيرة (عدد الغرف) (x_i)	التكرار (عدد العائلات) (n_i)	ت . ت . الصاعد $(n_i \uparrow)$	ت . ت . النازل $(n_i \downarrow)$
2	10	10	50
3	20	30	40
4	12	42	20
5	8	50	8
$\sum$	50	/	/

الشكل رقم (02-02): العرض البياني للتكرارات التجميعية الصاعدة



الشكل رقم (02-03): العرض البياني للتكرارات التجميعية النازلة



II-2-2-2-2. العروض البيانية في حالة متغير كمي مستمر (متصل):

إن العروض البيانية للمتغير الكمي المتصل هي المدرج التكراري، المضلع التكراري، التكرارات التجميعية الصاعدة والتكرارات التجميعية النازلة.

العرض البياني للمدرج والمضلع التكراريين:

إن **المدرج التكراري** هو عبارة عن أعمدة مستطيلات متلاصقة كل مستطيل قاعدته طول فئة وارتفاعه هو قيمة تكرار تلك الفئة، لكن لابد من التنويه إلى أنه إذا كانت أطوال الفئات متساوية يتم استخدام التكرارات المطلقة (n_i) أما إذا كانت أطوال الفئات غير متساوية فإنه يتم استخدام التكرارات المعدلة (n_i^*) ، أما **المضلع التكراري** فعادة ما يتم الاعتماد على المدرج التكراري لتمثيله ، فهو قطع مستقيمة متلاصقة تربط ما بين مراكز الفئات على قمم المستطيلات المتلاصقة، وبما أنه لدينا نوعين من المدرج التكراري، الأول خاص بفئات متساوية الأطوال والثاني خاص بفئات غير متساوية الأطوال فإننا نصادف نوعين من المضلع التكراري:

- المضلع التكراري في حالة أطوال الفئات متساوية يمثل بيانيا بتوصيل مراكز الفئات وفق التكرارات المطلقة المقابلة لها.
- المضلع التكراري في حالة أطوال الفئات غير متساوية يتم تحديد فئات متساوية الأطوال مع أصغر طول للفئات (L^*) ثم يتم تحديد مراكزها وبعدها يتم توصيل نقاط مراكز الفئات الجديدة مع التكرارات المقابلة لها.

وانطلاقا مما ذكرنا نميز حالتين عند رسم المدرج والمضلع التكراريين:

➢ **المدرج والمضلع التكراريين في حالة فئات متساوية الأطوال:** لما تكون الفئات متساوية الأطوال فإن نستخدم التكرارات المطلقة في رسم المدرج التكراري ، ولتوضيح ذلك نأخذ المثال التالي:

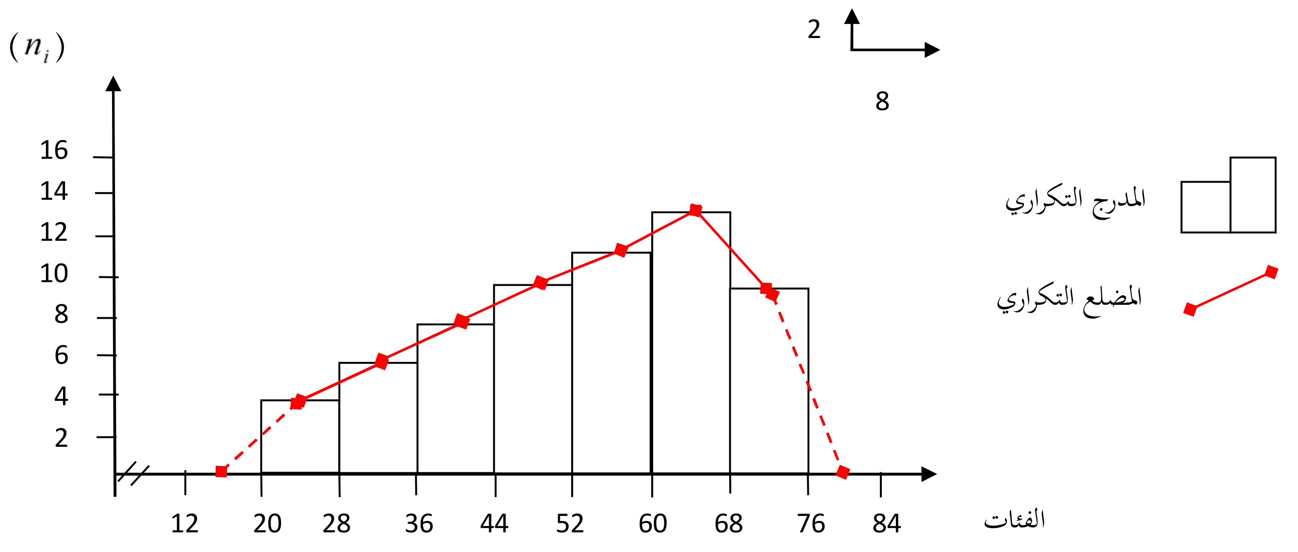
مثال (02–09): من معطيات المثال السابق (02-03) المبوبة في الجدول (02-08) السابق.

النفقات الشهرية (الفئات) X_i	عدد العائلات (التكرار) n_i	مركز الفئة C_i
$]28-20]$	4	24
$]36-28]$	6	32
$]44-36]$	8	40
$]52-44]$	10	48
$]60-52]$	12	56
$]68-60]$	14	64
$]76-68]$	10	72
$\sum$	64	-

المطلوب: مثل بيانات الجدول من خلال تمثيل بياني ملائم ؟

الحل: إن التمثيل البياني الملائم هو المدرج التكراري، وبما أن أطوال الفئات متساوية فإننا سنستخدم التكرارات المطلقة (n_i)، كما أن المعطيات تتوفر كذلك على مراكز الفئات وبالتالي سنستخدمها في رسم المضلع التكراري.

الشكل رقم (02-04): العرض البياني للمدرج والمضلع التكراريين في حالة فئات متساوية الأطوال



➢ **المدرج التكراري في حالة فئات غير متساوية الأطوال**: لما تكون الفئات غير متساوية الأطوال فإن نستخدم التكرارات المعدلة في رسم المدرج التكراري، ولتوضيح ذلك نأخذ المثال الموالي:

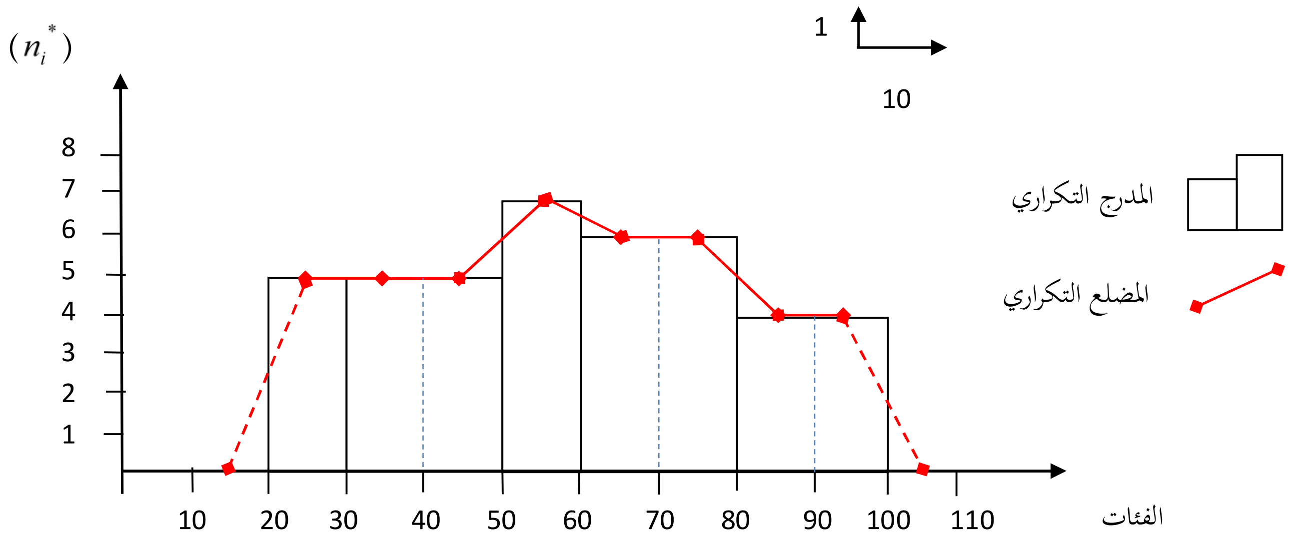
مثال (02-10): من معطيات المثال (02-07) السابق.

الفئات	التكرارات (n_i)	طول الفئات (L_i)	التكرارات المعدلة (n_i^*)
$]30-20]$	5	10	$n_1^* = \frac{n_1}{L_1} \times L^* = \frac{5}{10} \times 10 = 5$
$]50-30]$	10	20	$n_2^* = \frac{n_2}{L_2} \times L^* = \frac{10}{20} \times 10 = 5$
$]60-50]$	7	10	$n_3^* = \frac{n_3}{L_3} \times L^* = \frac{7}{10} \times 10 = 7$
$]80-60]$	12	20	$n_4^* = \frac{n_4}{L_4} \times L^* = \frac{12}{20} \times 10 = 6$
$]100-80]$	8	20	$n_5^* = \frac{n_5}{L_5} \times L^* = \frac{8}{20} \times 10 = 4$
المجموع	42	/	/

المطلوب: مثل بيانات الجدول من خلال تمثيل بياني ملائم ؟

الحل: نلاحظ أن أطوال الفئات غير متساوية وبالتالي من أجل رسم المدرج التكراري نستخدم التكرارات المعدلة الموجودة في الجدول السابق، أما المضلع التكراري فيتم تقسيم الفئات غير المتساوية الأطوال إلى فئات متساوية الطول مع أصغر طول فئة (طول الفئة المختار) كما يتم تكوين فئة جديدة قبل الفئة الأولى وأخرى بعد الفئة الأخيرة بدون تكرارات (تكراريهما معدوم)، ثم تحدد مراكز لكل الفئات وتوصل مراكزها في القمم بقطع مستقيمة متلاصقة، وبالتالي يصبح التمثيل البياني:

الشكل رقم (02-05): العرض البياني للمدرج والمضلع التكراريين في حالة فئات غير متساوية الأطوال



🖎 **العرض البياني للتكرارات المتجمعة الصاعدة**: يتم رسم منحنى التكرار التجميعي الصاعد للمتغيرة الكمية المستمرة (الفئات) من خلال إيصال نقاط الحدود العليا للفئات والتكرارات المتجمعة الصاعدة المقابلة لها، ثم يتم توصيل بداية التمثيل إلى الحد الأدنى للفئة الأولى.

🖎 **العرض البياني للتكرارات المتجمعة النازلة**: يتم رسم منحنى التكرار التجميعي الصاعد للمتغيرة الكمية المستمرة (الفئات) من خلال إيصال نقاط الحدود الدنيا للفئات والتكرارات المتجمعة النازلة المقابلة لها، ثم يتم توصيل نهاية التمثيل إلى الحد الأعلى للفئة الأخيرة.

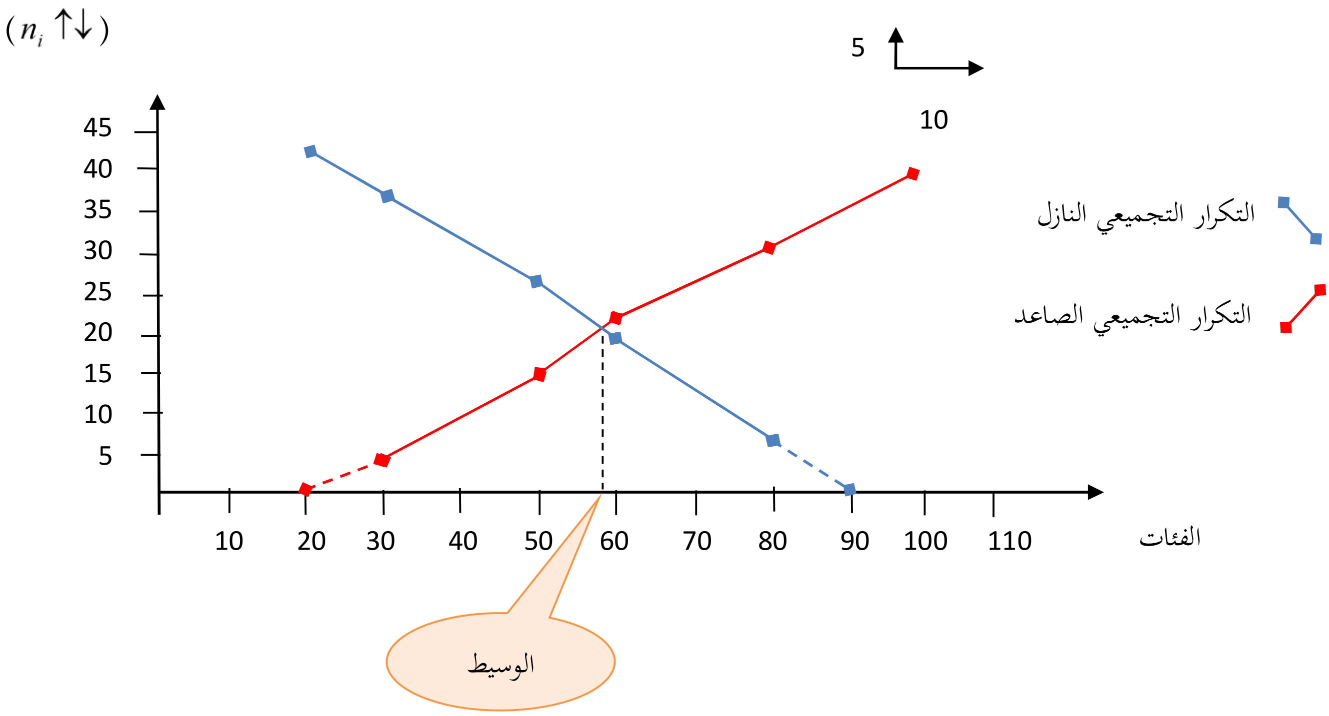
لتوضيح العرض البياني للتكرارات التجميعية الصاعدة والنازلة للمتغيرة الكمية المستمرة نستخدم معطيات أحد الأمثلة السابقة.

مثال (02-11): من معطيات المثال (02-07) المبينة في الجدول الموالي.

الفئات	التكرارات (n_i)	طول الفئات (L_i)	ت . ت الصاعد $(n_i \uparrow)$	ت.ت.النازل $(n_i \downarrow)$
$]30-20]$	5	10	5	42
$]50-30]$	10	20	15	37
$]60-50]$	7	10	22	27
$]80-60]$	12	20	34	20
$]100-80]$	8	20	42	8
المجموع	42	/	/	/

المطلوب: مثل بيانيا التكرارات التجميعية الصاعدة والنازلة ؟

الشكل رقم (02–06): العرض البياني للتكرارات التجميعية الصاعدة والنازلة.



II–2–2–2–3. العروض البيانية في حالة متغير كيفي (نوعي):

توجد ثلاثة عروض بيانية للمتغير الكيفي هي العرض الدائري، العمود المجزأ والأعمدة المستطيلة.

- **العرض البياني الدائري:** هو عبارة عن دائرة (360°) مقسمة إلى عدة أجزاء (زوايا) كل جزء (زاوية) يتناسب مع التكرارات المقابلة لكل صفة من الصفات المدروسة، ويتم حساب كل زاوية وفق القاعدة التالية :

$$\text{الزاوية المركزية} = \frac{\text{تكرار الصفة}}{\text{مجموع التكرارات}} \times 360° = \text{التكرار النسبي} \times 360°$$

$$\text{الزاوية المركزية} = f_i \times 360°$$

- **العرض البياني للعمود المجزأ:** هو عبارة عن عمود مستطيل مقسم إلى عدة أجزاء، يستخدم في تجزئته التكرار النسبي المئوي ($f_i\%$)، حيث ارتفاع المستطيل يعادل مجموع التكرارات النسبية المئوية (100 %)، وكل جزء يمثل صفة وإرتفاعه يعادل التكرار النسبي المئوي لتلك الصفة.
- **العرض البياني للأعمدة المستطيلة:** هي عبارة عن أعمدة مستطيلات غير متلاصقة، وتكون متباعدة بمسافات ثابتة، حيث كل عمود مستطيل يمثل صفة وإرتفاعه يتناسب مع التكرار المقابلة له.

لتوضيح كيفية تمثيل العروض البيانية الخاصة بالمتغير النوعي نستخدم المثال الموالي:

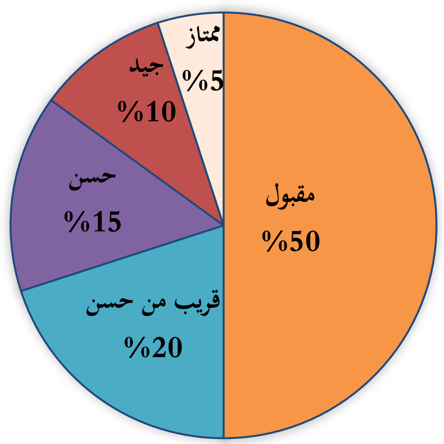
مثال (02–12): تمثل بيانات المبوبة في الجدول الموالي تقديرات النجاح في شهادة البكالوريا لمجموعة من تلاميذ مدرسة.

المتغير (تقديرات النجاح)	التكرار (عدد التلاميذ)	التكرار النسبي f_i	التكرار النسبي المئوي f_i %	الزاوية المركزية (°360 x f_i)
مقبول	50	0.5	% 50	° 180
قريب من حسن	20	0.2	% 20	° 72
حسن	15	0.15	% 15	° 54
جيد	10	0.10	% 10	° 36
ممتاز	5	0.05	% 5	° 18
المجموع	100	1	%100	° 360

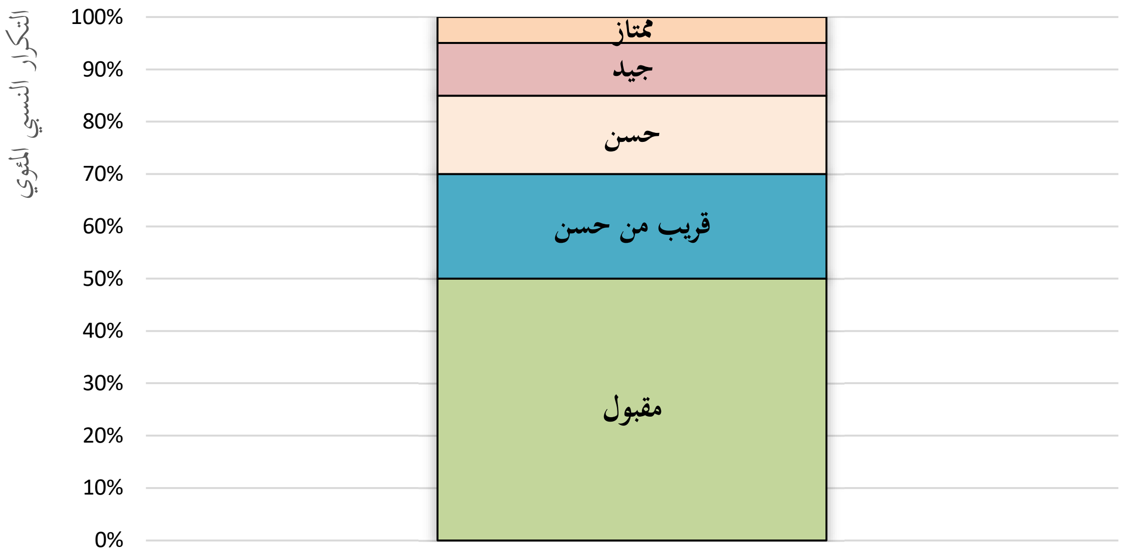
المطلوب: أرسم بيانات الجدول بإستخدام العروض البيانية ؟

الحل: هناك ثلاثة عروض بيانية هي العرض الدائري، العرض البياني بواسطة العمود المجزأ والعرض البياني بواسطة الأعمدة المستطيلة.

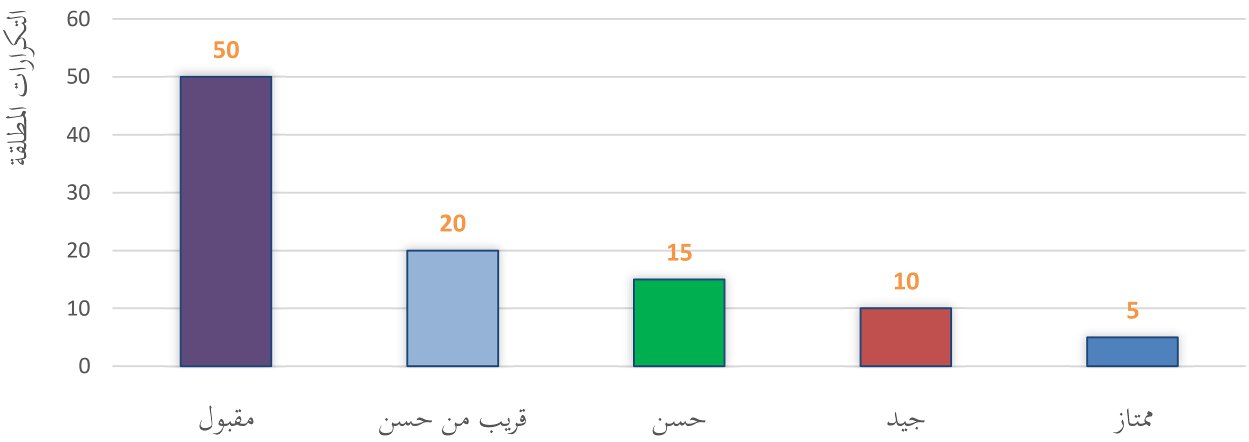
➢ **الشكل رقم (02–07):** العرض الدائري



➢ **الشكل رقم (02–08)**: العرض البياني بواسطة العمود المجزأ



➢ **الشكل رقم (02–09)**: العرض البياني بواسطة الأعمدة المستطيلة



المحور الثالث:

مقاييس النزعة المركزية

المحور الثالث : مقاييس النزعة المركزية

III-1. تمهيد:

بعدما قدمنا في المحور السابق العرض الجدولي والبياني للبيانات الإحصائية، ولكي يكتمل التحليل المثالي لهذه البيانات لابد من عرض مقاييس مساعدة تحدد مجموعة من المقاسات أبرزها تباعد القيم عن بعضها البعض، أو تباعد القيم عن قيمة محددة، أو اختلاف البيانات عن بيانات معيارية (مثالية) ...إلخ، ولكل من هذه المقاييس استخدام خاص بها.

إن مقاييس النزعة المركزية هي أول مقاييس يبدأ بها الباحث، لأنها تعطي له نظرة عن تموقع قيمه (بياناته) حول قيمتها المتوسطة أو المتمركزة، وقبل عرض هذه المقاييس لابد من عرض المفهوم الاحصائي لهذه المقاييس.

III-2. تعريف مقياس النزعة المركزية:

هو عبارة عن قيمة متوسطة تتمركز حولها بيانات الظاهرة، وهي كذلك تعبر أو تمثل جميع البيانات، وفي نفس الوقت تعطي فكرة واضحة عن الظاهرة المدروسة.

III-3. أنواع مقاييس النزعة المركزية:

تنقسم مقاييس النزعة المركزية إلى عدة أنواع أكثرها إستخداما: المتوسط الحسابي، الوسيط والمنوال، لكن توجد أنواع أخرى من بينها أشباه المتوسط الحسابي وأشباه الوسيط.

III-3-1. المتوسط الحسابي (The Arithmetic Mean):

إن المتوسط الحسابي هو أشهر مقياس نزعة مركزية وأكثرها إستخداما في الإحصاء، ويرمز له بالرمز $\overline{X}$.

بالنظر إلى وجود نوعين من البيانات وهما البيانات الخام أو الأولية (غير المبوبة) والبيانات المنظمة (المبوبة) فإن طريقة حساب قيمة المتوسط الحسابي تختلف حسب نوعية البيانات.

III-3-1-1. حساب المتوسط الحسابي من البيانات الأولية: [1]

الطريقة المباشرة:

إذا كانت قيم الظاهرة المدروسة هي : x_1 ; x_2 ; .... ; x_n والتي عددها n فإن المتوسط الحسابي لهذه القيم هو مجموع هذه القيم مقسوما على عددها (n) ، ويعطى بالصيغة التالية:

$$\overline{X} = \frac{x_1 + x_2 + \dots\dots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

حيث : n هو عدد القيم

مثال (3-01): لتكن القيم التالية.

2 ، 4 ، 8 ، 9 ، 10 ، 8 ، 10 ، 9 ، 10 ، 7 ، 5 ، 8

المطلوب : حساب متوسط القيم ؟

[1] - MURRAY R. SPIEGEL and LARRY J. STEPHENS ; 2008 ; Theory and Problems of STATISTICS ; 4 th Edition ; Schaum's Outline Series ; PP : 62 , 63.

الحل:

$$\overline{X} = \frac{2+4+8+9+10+8+10+9+10+7+5+8}{12} = \frac{90}{12} = 7.5$$

مثال (3–02): لتكن البيانات الموالية تمثل عدد أطفال مجموعة من العائلات

الجدول رقم (3–01): توزيع العائلات حسب عدد أطفالها.

رقم العائلة	1	2	3	4	5	6	7	8	9
عدد أطفالها	5	3	6	2	1	4	4	6	2

المطلوب : أحسب متوسط عدد أطفال العائلات

الحل:

$$\overline{X} = \frac{5+3+6+2+1+4+4+6+2}{9} = \frac{33}{9} = 3.66 \cong 4$$

ملاحظة : تم تقريب العدد 3.66 إلى 4 لأن نوع المتغير منفصل (متقطع) وهو يمثل عدد الأطفال (عدد صحيح) على عكس المثال السابق الذي لم يتم تحديد نوع المتغير (قيل بأنها قيم).

طريقة الوسط الفرضي:

إذا كانت قيم الظاهرة المدروسة هي : x_1 ; x_2 ; .... ; x_n والتي عددها n فإن المتوسط الحسابي لهذه القيم بواسطة طريقة الوسط الفرضي هو الوسط الفرضي (x_0) مضاف إليه مجموع فروقات القيم عن هذا الوسط الفرضي مقسوم على عدد القيم وفق القاعدة التالية :

$$\overline{X} = x_0 + \frac{\sum_{i=1}^{n}(x_i - x_0)}{n}$$

حيث :

x_i : قيم الظاهرة المدروسة.

x_0 : الوسط الفرضي ، حيث من الأفضل أن يكون أحد قيم الظاهرة المدروسة.

n : عدد القيم.

مثال (3–03): لتكن قيم المثال السابق الموالية.

2 ، 4 ، 8 ، 9 ، 10 ، 8 ، 10 ، 9 ، 10 ، 7 ، 5 ، 8

المطلوب : أحسب المتوسط الحسابي للقيم بواسطة طريقة الوسط الفرضي ؟

الحل :

لنفترض أن الوسط الفرضي هو $x_0 = 8$

$$\overline{X} = x_0 + \frac{\sum_{i=1}^{n}(x_i - x_0)}{n}$$

$$= 8 + \frac{(2-8)+(4-8)+(8-8)+(9-8)+(10-8)+(8-8)+(10-8)+(9-8)+(10-8)+(7-8)+(5-8)+(8-8)}{12}$$

$$= 8 + \frac{-6-4+0+1+2+0+2+1+2-1-3+0}{12} = 8 - \frac{6}{12}$$

$$\overline{X} = 7.5$$

III-3-1-2. حساب المتوسط الحسابي من البيانات المبوبة:

نصادف نوعين من البيانات المبوبة، بيانات خاصة بمتغير كمي منفصل (متقطع) ذات قيم نقطية، وبيانات متعلقة بمتغير كمي مستمر (متصل) ذات فئات.

III-3-1-2-1. حساب المتوسط الحسابي لمتغير كمي منفصل (منقطع):

يوجد طريقتين لحساب المتوسط الحسابي في هذه الحالة، طريقة مباشرة وطريقة وسط فرضي.

الطريقة المباشرة:

إذا كانت قيم الظاهرة المدروسة هي : $x_1\ ;\ x_2\ ;\ ....\ ;\ x_k$ ، وكانت $n_1\ ;\ n_2\ ;\ ....\ ;\ n_k$ تمثل التكرارات المقابلة لها، فإن المتوسط الحسابي لهذه القيم بواسطة الطريقة المباشرة يحسب وفق القاعدة التالية:

$$\overline{X} = \frac{x_1 n_1 + x_2 n_2 + ... + x_k n_k}{n_1 + n_2 + ... + n_k}$$

$$\overline{X} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i}$$

مثال (3-04): لتكن البيانات الإحصائية الموالية تمثل عدد أيام التغيب لدى مجموعة عمال مؤسسة ما خلال فترة شهر.

الجدول رقم (3-02): توزيع العمال حسب عدد أيام التغيب.

عدد أيام الغياب X_i	0	1	2	3	4	5	$\sum$
عدد العمال n_i	10	8	4	2	3	3	30

المطلوب: إيجاد متوسط عدد أيام التغيب لدى مجموعة عمال المؤسسة ؟

الحل: لدينا:

$$\overline{X} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i}$$

$$\overline{X} = \frac{(0\times10)+(1\times8)+(2\times4)+(3\times2)+(4\times3)+(5\times3)}{30}$$

$$\overline{X} = \frac{49}{30} = 1.63$$

$$\overline{X} = 2$$

إذن متوسط عدد أيام التغيب لدى عمال المؤسسة هو يومان.

ملاحظة : يمكن إستخدام جدول التوزيع التكراري في حساب المتوسط الحسابي وهذا بإضافة عمود ثالث إلى الجدول يخصص لحاصل ضرب قيمة المتغير الإحصائي في التكرار المقابل لهذه القيمة $(x_i \times n_i)$ كما هو موضح في الجدول الموالي.

الجدول رقم (03-3): طريقة حساب المتوسط الحسابي بإستخدام جدول التوزيع التكراري.

عدد أيام التغيب X_i	عدد العمال n_i	$x_i n_i$
0	10	0
1	8	8
2	4	8
3	2	6
4	3	12
5	3	15
المجموع	30	49

إذن متوسط عدد أيام التغيب لدى عمال المؤسسة هو:

$$\overline{X} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i} = \frac{49}{30} = 1,63 \approx 2$$

طريقة الوسط الفرضي:

يتم حساب المتوسط الحسابي وفق نفس الطريقة المستخدمة في البيانات غير المبوبة، فإذا كانت قيم الظاهرة المدروسة هي : x_1 ; x_2 ; .... ; x_k ، وكانت n_1 ; n_2 ; .... ; n_k تمثل التكرارات المقابلة لها، فإن المتوسط الحسابي لهذه القيم بواسطة طريقة الوسط الفرضي يحسب وفق القاعدة التالية:

$$\overline{X} = x_0 + \frac{\sum_{i=1}^{k} (x_i - x_0)\ n_i}{\sum_{i=1}^{k} n_i}$$

حيث :

x_i : قيم الظاهرة المدروسة (القيم النقطية).

x_0 : الوسط الفرضي، ومن الأفضل أن يكون أحد قيم الظاهرة المدروسة.

n_i : تكرارات القيم.

مثال (3–05): لتكن البيانات الإحصائية الخاصة بالمثال السابق الموالية، والتي كانت تمثل عدد أيام التغيب لدى مجموعة عمال مؤسسة ما خلال فترة شهر.

الجدول رقم (3–04): معطيات المثال السابق والتي تمثل توزيع العمال حسب عدد أيام التغيب.

$\sum$	5	4	3	2	1	0	عدد أيام الغياب X_i
30	3	3	2	4	8	10	عدد العمال n_i

المطلوب: إيجاد متوسط عدد أيام التغيب لدى مجموعة عمال المؤسسة باستخدام طريقة الوسط الفرضي؟

الحل: بالاعتماد على جدول التوزيع التكراري يتم حساب المتوسط الحسابي.

الجدول رقم (3–05): حساب المتوسط الحسابي بواسطة طريقة الوسط الفرضي $(x_0 = 3)$

$(x_i - x_0)\ n_i$	$x_i - x_0$	عدد العمال n_i	عدد أيام التغيب X_i
-30	-3	10	0
-16	-2	8	1
-4	-1	4	2
0	0	2	3
3	1	3	4
6	2	3	5
-41	/	30	المجموع

إذن المتوسط الحسابي باستخدام طريقة الوسط الفرضي هو:

$$\overline{X} = x_0 + \frac{\sum_{i=1}^{k} (x_i - x_0)\ n_i}{\sum_{i=1}^{k} n_i} = 3 - \frac{41}{30} = 1.63 \cong 2$$

$$\overline{X} = 2$$

III–3–1–2–2. حساب المتوسط الحسابي لمتغير كمي مستمر (متصل):

يوجد طريقتين لحساب المتوسط الحسابي في حالة متغير كمي مستمر، طريقة مباشرة وطريقة وسط فرضي.

الطريقة المباشرة:

يتم حساب المتوسط الحسابي في حالة متغير كمي مستمر اعتمادا على مراكز الفئات (C_i)، فإذا كانت مراكز الفئات هي : $C_1\ ;\ C_2\ ;\ ....\ ;\ C_k$ ، وكانت $n_1\ ;\ n_2\ ;\ ....\ ;\ n_k$ تمثل التكرارات المقابلة لها، فإن المتوسط الحسابي لهذه الفئات بواسطة الطريقة المباشرة يحسب وفق القاعدة التالية:

$$\overline{X} = \frac{C_1 n_1 + C_2 n_2 + ... + C_k n_k}{n_1 + n_2 + ... + n_k}$$

$$\overline{X} = \frac{\sum_{i=1}^{k} C_i n_i}{\sum_{i=1}^{k} n_i}$$

حيث :

C_i : مركز الفئة i ، و i=1 ; 2 ;..... ; k

n_i : تكرار الفئة i

k : عدد الفئات

مثال (3-06): لتكن البيانات الإحصائية الموالية التي تمثل الدخل الشهري (الوحدة: 10^3دج) لمجموعة من الأفراد.

الجدول رقم (3-06): توزيع الأفراد حسب دخلهم الشهري.

فئات الدخل الشهري x_i	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: إيجاد متوسط دخل الأفراد باستخدام الطريقة المباشرة ؟

الحل: إعتمادا على جدول التوزيع التكراري يتم حساب المتوسط الحسابي وهذا بعد استخراج مراكز الفئات.

الجدول رقم (3-07): حساب المتوسط الحسابي اعتمادا على جدول التوزيع التكراري.

فئات الدخل الشهري X_i	عدد الأفراد n_i	مراكز الفئات c_i	$c_i n_i$
$]20-10]$	5	15	75
$]30-20]$	8	25	200
$]40-30]$	12	35	420
$]50-40]$	25	45	1125
$]60-50]$	10	55	550
المجموع	60	/	2370

إذن المتوسط الحسابي لدخل الأفراد بواسطة الطريقة المباشرة هو :

$$\overline{X} = \frac{\sum_{i=1}^{k} C_i n_i}{\sum_{i=1}^{k} n_i} = \frac{2370}{60} = 39,5$$

$$\overline{X} = 39.5 \times 10^3 \ DA$$

طريقة الوسط الفرضي:

يتم حساب المتوسط الحسابي لبيانات متغير كمي متصل وفق نفس الطريقة المستخدمة في البيانات غير المبوبة، فإذا كانت مراكز الفئات هي : $C_1 ; C_2 ; .... ; C_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، فإن المتوسط الحسابي لهذه الفئات بواسطة طريقة الوسط الفرضي يحسب وفق القاعدة التالية:

$$\overline{X} = C_0 + \frac{\sum_{i=1}^{k} (C_i - C_0)\ n_i}{\sum_{i=1}^{k} n_i}$$

حيث :

C_i : مركز الفئة i ، و $i=1 ; 2 ;..... ; k$

C_0 : مركز الفئة الفرضي، ومن الأفضل أن يكون أحد مراكز فئات المتغيرة المدروسة.

n_i : تكرار الفئة i.

k : عدد الفئات

مثال (3–07): لتكن البيانات الإحصائية الموالية الخاصة بالمثال السابق، والتي كانت تمثل الدخل الشهري (الوحدة: 10^3دج) لمجموعة من الأفراد.

الجدول رقم (3–08): بيانات المثال السابق والتي تمثل توزيع الأفراد حسب دخلهم الشهري.

فئات الدخل الشهري x_i	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: إيجاد متوسط دخل الأفراد باستخدام طريقة الوسط الفرضي ؟

الحل: إعتمادا على جدول التوزيع التكراري يتم حساب المتوسط الحسابي وهذا بعد استخراج مراكز الفئات.

الجدول رقم (3–09): حساب المتوسط الحسابي اعتمادا على جدول التوزيع التكراري مع افتراض أن $C_0 = 35$

فئات الدخل الشهري X_i	عدد الأفراد n_i	مراكز الفئات c_i	$C_i - C_0$	$(C_i - C_0)\ n_i$
$]20-10]$	5	15	-20	-100
$]30-20]$	8	25	-10	-80
$]40-30]$	12	35	0	0
$]50-40]$	25	45	10	250
$]60-50]$	10	55	20	200
المجموع	60	/	/	270

إذن المتوسط الحسابي لدخل الأفراد بواسطة طريقة الوسط الفرضي هو:

$$\overline{X} = C_0 + \frac{\sum_{i=1}^{k} (C_i - C_0)\ n_i}{\sum_{i=1}^{k} n_i} = 35 + \frac{270}{60}$$

$$\overline{X} = 39.5 \times 10^3\ DA$$

III-3-1-3. حساب المتوسط الحسابي كمتوسط حسابي مرجح لمتوسطات حسابية:

إذا كانت مجموعة بيانات أولى عددها n_1 ومتوسطها الحسابي $\overline{X_1}$ ، وكانت مجموعة بيانات ثانية عددها n_2 ومتوسطها الحسابي $\overline{X_2}$ ، وجرى دمج المجموعتين فإن المتوسط الحسابي للمجموعتين يحسب إنطلاقا من عدد بيانات المجموعتين ومتوسطيهما الحسابيين وفق القاعدة التالية :

$$\overline{X} = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2}$$

مثال (3-08): تمثل البيانات الموالية عدد الأطفال لدى مجموعتين من العائلات في عمارتين مختلفين لحي ما.

عدد أطفال عائلات العمارة الأولى هو: 5 ، 3 ، 4 ، 3 ، 6 ، 2 ، 4 ، 5

عدد أطفال عائلات العمارة الثانية هو: 2 ، 4 ، 3 ، 1 ، 4 ، 3 ، 2 ، 4 ، 4 ، 3

المطلوب: إيجاد متوسط عدد الأطفال في كل عمارة، ثم متوسط عدد الأطفال في العمارتين معا ؟

الحل:

- حساب متوسط عدد أطفال عائلات العمارة الأولى:

$$\overline{X_1} = \frac{\sum_{i=1}^{8} X_i}{n_1} = \frac{5+3+4+3+6+2+4+5}{8} = \frac{32}{8} = 4$$

- حساب متوسط عدد أطفال عائلات العمارة الثانية:

$$\overline{X_2} = \frac{\sum_{i=1}^{10} X_i}{n_2} = \frac{2+4+3+1+4+3+2+4+4+3}{10} = \frac{30}{10} = 3$$

- حساب متوسط عدد أطفال العمارتين معا: يمكن إستخدام البيانات الخاصة بالعمارتين أو إنطلاقا مما قمنا بحسابه، أي المتوسطين الحسابيين وعدد بيانات كل عمارة.

ط01 : نعتمد على جميع البيانات

$$\overline{X} = \frac{\sum_{i=1}^{18} X_i}{n} = \frac{5+3+4+3+6+2+4+5+2+4+3+1+4+3+2+4+4+3}{18} = \frac{62}{8} = 3.44 \approx 3$$

ط02 : نعتمد على المتوسطات الحسابية

$$\overline{X} = \frac{n_1\overline{x_1} + n_2\overline{x_2}}{n_1 + n_2} = \frac{8\times4+10\times3}{8+10} = \frac{32+30}{18} = \frac{62}{18} = 3.44 \approx 3$$

إذن متوسط عدد أطفال عمارتي الحي هو 3 أطفال.

III-3-1-4. خواص ومميزات المتوسط الحسابي: للمتوسط الحسابي مجموعة من المحاسن والمساوئ نلخصها فيما يلي: [1]

من محاسن المتوسط الحسابي أنه:

- يعتبر أكثر إستعمالا ووضوحا من جميع المتوسطات الأخرى التي سندرسها لاحقا؛
- بالرغم ما يستغرق من وقت لحسابه، فإن حسابه يبقى سهل وغير معقد؛
- عندما يكون للوسط الحسابي قيمة ممثلة، فإنه يستعمل كضابط أو معيار تقارن به المفردات الأخرى في التوزيع.

من مساوئ المتوسط الحسابي أنه:

- يتأثر بوجود القيم المتطرفة؛
- عندما تكون فئات التوزيع مفتوحة من غحدى الطرفين يصعب حسابه.

III-3-2. أشباه المتوسط الحسابي: في بعض الظواهر المتوسط الحسابي لا يصف وصف سليم بيانات تلك الظواهر، ولا يعطي أي فكرة صحيحة عنها، وبالتالي توجد متوسطات أخرى قد تعبر عن البيانات بصورة أفضل، ومن هذه المتوسطات: المتوسط الهندسي، المتوسط التوافقي، المتوسط التربيعي.

III-3-2-1. المتوسط الهندسي (The Geometric Mean):

يستخدم المتوسط الهندسي عندما تكون قيم الظاهرة المدروسة عبارة عن نسب أو معدلات مثل معدلا نمو الظواهر، ويرمز له بـــ **G**.

بما أنه يوجد نوعين من البيانات، بيانات خام وأخرى مبوبة، سنعرض قاعدة المتوسط الهندسي حسب نوع البيانات.

III-3-2-1-1. حساب المتوسط الهندسي للبيانات الأولية (غير المبوبة):

إن المتوسط الهندسي لمجموعة من القيم (المعدلات أو النسب) x_1 ; x_2 ; .... ; x_n والتي عددها **n** هو عبارة عن الجذر النوني لجداءات تلك القيم[2]، والقاعدة تكتب كما يلي:

$$G = \sqrt[n]{x_1 \times x_2 \times ..... \times x_n}$$

مثال (3-09): تمثل البيانات الموالية على الترتيب معدلات النمو الاقتصادي للدول المغاربية: ليبيا، تونس، الجزائر، المغرب وموريطانيا : 2% ، 3% ، 4% ، 4% ، 1.5%

المطلوب: أحسب متوسط معدل نمو المنطقة المغاربية ؟

[1] - محمد بوهزة، مرجع سابق ، ص ص : 59 ، 60.

[2] - محمد راتول، مرجع سابق، ص 119.

الحل: بما أن المعطيات متعلقة بمعدل نمو فالمتوسط الحسابي الملائم هو المتوسط الهندسي.

$$G=\sqrt[n]{x_1\times x_2\times.....\times x_n}=\sqrt[5]{0.02\times0.03\times0.04\times0.04\times0.015}=0.027$$

ملاحظة:

من أجل تسهيل العملية الحسابية في حالة عدم معرفة إجراء العمليات الحسابية بالجذر النوني، أو الصغر الكبير أثناء ضرب هذه البيانات يتم حساب المتوسط الهندسي باستخدام لوغاريتم البيانات، ويكون ذلك كما يلي:

$$G=\sqrt[n]{x_1\times x_2\times.....\times x_n}=(x_1\times x_2\times.....\times x_n)^{1/n}$$

ندخل اللوغاريتم على الطرفين فنتحصل على:

$$\ln G=\ln(x_1\times x_2\times......\times x_n)^{1/n}$$

ومن خواص اللوغاريتم نحصل على ما يلي:

$$\ln G=\frac{1}{n}\ln(x_1\times x_2\times......\times x_n)$$

$$\ln G=\frac{1}{n}(\ln x_1+\ln x_2+......+\ln x_n)$$

$$\ln G=\frac{1}{n}(\sum_{i=1}^{n}\ln x_i)$$

وللتخلص من اللوغاريتم يتم إدخال الدالة الأسية Exponentielle كالتالي:

$$e^{\ln G}=e^{\frac{1}{n}(\sum_{i=1}^{n}\ln x_i)}$$

ومن خواص الدالة الأسية تصبح علاقة المتوسط الهندسي كما يلي:

$$G=e^{\frac{1}{n}\sum_{i=1}^{n}\ln x_i}$$

مثال (3–10): من معطيات المثال السابق (3-09) يمكننا حساب المتوسط الهندسي بطريقة اللوغاريتم ويكون ذلك كما يلي:

$$G=e^{\frac{1}{n}\sum_{i=1}^{n}\ln x_i}=e^{\frac{1}{5}(\ln0.02+\ln0.03+\ln0.04+\ln0.04+\ln0.015)}$$

$$G=e^{\frac{1}{n}\sum_{i=1}^{n}\ln x_i}=e^{\frac{1}{5}(\ln0.02+\ln0.03+\ln0.04+\ln0.04+\ln0.015)}$$

$$G=e^{-3.611}=0.027$$

III-3-2-1-2. حساب المتوسط الهندسي للبيانات المبوبة:

إن المتوسط الهندسي عادة ما يستخدم عندما تكون البيانات عبارة عن معدلات نمو ظاهرة ، وبالتالي يمكن القول أنه يستخدم في البيانات المبوبة للمتغير الكمي المستمر، ولا يمكن أن يحسب عند المتغير الكمي المتقطع لأنه قيم نقطية (صحيحة).

إذن إذا كانت مراكز الفئات هي : $C_1 ; C_2 ; .... ; C_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، فإن المتوسط الهندسي لهذه الفئات يحسب وفق القاعدة التالية:

$$G = \sqrt[N]{C_1^{\,n_1} \times C_2^{\,n_2} \times ..... \times C_k^{\,n_k}}$$

حيث:

$N = \sum_{i=1}^{k} n_i$: مجموع التكرارات.

كما يمكن إستخدام القاعدة السابقة وهي قاعدة اللوغاريتم، ويصبح المتوسط الهندسي يحسب كما يلي:

$$G = e^{\frac{1}{N}\sum_{i=1}^{k} n_i \ln C_i}$$

حيث : $e = 2,718$ وهي الدالة الأسية

$\ln$ اللوغاريتم النيبيري

مثال (3-10): تمثل البيانات الموالية معدلات النمو السكاني لـ 30 دولة

الجدول رقم (3-10): توزيع البلدان حسب معدلات نمو سكانها

فئات معدلات نمو السكان	]0.03 – 0.01]	]0.05 – 0.03]	]0.07 – 0.05]	]0.09 – 0.07]	]0.11 – 0.09]
عدد الدول (n_i)	10	8	6	4	2

المطلوب: أحسب متوسط معدلات النمو السكاني للبلدان ؟

الحل: بما أن البيانات متعلقة بمعدلات نمو ظاهرة فإن المتوسط الحسابي الملائم هو المتوسط الهندسي.

بإستخدام جدول التوزيع التكراري نقوم بحساب المتوسط الهندسي وفق الطريقتين السابقتين.

الجدول رقم (3-11): حساب المتوسط الهندسي إعتمادا على جدول التوزيع التكراري

الفئات	n_i	C_i	$C_i^{\,n_i}$	$\ln C_i$	$n_i \ln C_i$
]0.03 – 0.01]	10	0.02	1.024e-17	-3.912	-39.12
]0.05 – 0.03]	8	0.04	6.5536e-12	-3.218	-25.75
]0.07 – 0.05]	6	0.06	4.6656e-8	-2.813	-16.88
]0.09 – 0.07]	4	0.08	4.096e-5	-2.525	-10.10
]0.11 – 0.09]	2	0.1	0.01	-2.302	-4.60
المجموع	30	/	$\prod_{i=1}^{5} C_i^{\,n_i} = 1.28247e-42$	/	-96.45

ط01 :

$$G=\sqrt[N]{C_1^{\ n_1}\times C_2^{\ n_2}\times.....\times C_k^{\ n_k}}$$

$$G=\sqrt[30]{0.02^{10}\times0.04^{8}\times0.06^{6}\times0.08^{4}\times0.1^{2}}$$

$$G=\sqrt[30]{1.024e-17\times6.5536e-12\times4.6656e-8\times4.096e-5\times0.01}$$

$$G=\sqrt[30]{1.28247e-42}=0.0402$$

ط02 :

$$G=e^{\frac{1}{N}\sum_{i=1}^{k}n_i\ln x_i}=e^{\frac{1}{30}(-96.45)}$$

$$G=e^{-3.215}=0.0401$$

مثال (3-11): تمثل البيانات الموالية تطور الدخل السنوي لأحد الأفراد خلال الفترة 2010-2017

الجدول رقم (3-12): توزيع دخل الفرد خلال الفترة (2010-2017) (الوحدة : 10^3 دج)

السنة	2010	2011	2012	2013	2014	2015	2016	2017
دخل الفرد	32	38	42	44	48	50	52	58

المطلوب: أحسب متوسط معدل نمو دخل الفرد ؟

الحل: قبل حساب متوسط معدل أو نسبة نمو دخل الفرد لابد من حساب معدل النمو خلال كل سنة وفق القاعدة التالية:

$$t_i=\frac{X_i-X_{i-1}}{X_{i-1}}\times100$$

إذن معدلات نمو دخل الفرد تكون كما يلي:

الجدول رقم (3-13): حساب معدلات نمو دخل الفرد السنوية

السنة	2010	2011	2012	2013	2014	2015	2016	2017
دخل الفرد	32	38	42	44	48	50	52	58
معدل نمو الدخل	-	0.187	0.105	0.047	0.090	0.041	0.04	0.115

إذن متوسط معدلات نمو دخل الفرد هو :

$$G=\sqrt[n]{x_1\times x_2\times.....\times x_n}=\sqrt[7]{0.187\times0.105\times0.047\times0.090\times0.041\times0.04\times0.115}$$

$$G=0.0767=7.67\%$$

متوسط نمو دخل الفرد هو 7.67%.

ملاحظة : كما أشرنا إليه سابقا عادة يتم إستخدام المتوسط الهندسي عند حساب معدلات نمو الظواهر(المتغير الكمي المستمر)، لكن إذا طلب حسابه من غير ذلك (المتغير الكمي المنقطع) فالصيغة الرياضية الخاصة بهذا الأخير هي كالتالي:

$$G = \sqrt[N]{{x_1}^{n_1} \times {x_2}^{n_2} \times ..... \times {x_k}^{n_k}}$$

أو

$$G = e^{\frac{1}{N}\sum_{i=1}^{k} n_i \ln x_i}$$

حيث: $N = \sum_{i=1}^{k} n_i$ مجموع التكرارات

$x_1 ; x_2 ; .... ; x_k$: قيم المتغير الكمي المنقطع

$e = 2,718$ وهي الدالة الأسية

$\ln$ اللوغاريتم النيبيري

III-3-2-1-3. **خواص المتوسط الهندسي**: من أهم خواص المتوسط الهندسي أنه:[1]

- نادر الاستخدام لصعوبة حسابه، ويستخدم في البيانات التي تشكل متتالية هندسية كبيانات تطور عدد السكان، او المبالغ المستثمرة وفق فائدة مركبة؛
- حاصل ضرب مجموعة من القيم لا يتغير إذا استبدلت كل قيمة من هذه القيم بالمتوسط الهندسي؛
- يكون المتوسط الهندسي دائما أقل من المتوسط الحسابي ولا يتساوى معه إلا إذا كانت جميع قيم الظاهرة متساوية؛
- المتوسط الهندسي يأخذ بعين الاعتبار جميع مفردات القيم؛
- لا يتأثر بالقيم المتطرفة؛
- ومن عيوبه انه لا يمكن حسابه في حالة بيانات التوزيعات التكرارية المفتوحة أو التي يكون جداءها سالبا.

III-3-2-2. المتوسط التوافقي (Harmonic Mean):

إذا كانت بيانات الظاهرة متعلقة بمعدلات السرعة أو متوسط الأسعار أو متوسط الكثافة السكانية فإن المتوسط الملائم والأكثر وصفا للظاهرة هو المتوسط التوافقي الذي يرمز له بالرمز H، وبما أنه يوجد نوعين من البيانات وهما البيانات المبوبة والبيانات غير المبوبة فإننا سنعرض القاعدة الرياضية للمتوسط التوافقي الموافقة لكل نوع.

III-3-2-2-1. حساب المتوسط التوافقي للبيانات الأولية (غير المبوبة):

إن المتوسط التوافقي لمجموعة من القيم (معدلات السرعة أو متوسط الأسعار او متوسط الكثافة السكانية) $x_1 ; x_2 ; .... ; x_n$ والتي عددها **n** هو عبارة عن عدد القيم على مجموع مقاليب تلك القيم، والصيغة الرياضية هي:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}}$$

[1] - محمد راتول، مرجع سابق، ص ص : 122 ، 123.

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

مثال (3–12): تمثل البيانات الموالية السرعة التي سار بها دراج أثناء قطعه 4 مسافات متساوية، حيث قطع المسافة الأولى بسرعة 20 كلم/سا ، ثم قطع المسافة الثانية بسرعة 60 كلم/سا ، ثم قطع المسافة الثالثة بسرعة 40 كلم/سا، ثم قطع المسافة الرابعة بسرعة 30 كلم/سا.

المطلوب: ما هو متوسط سرعة هذا الدراج؟

الحل: بما البيانات متعلقة بمعدلات السرعة فإن المتوسط الملائم هو المتوسط التوافقي، وكذلك المسافات متساوية، أي التكرارات متساوية فإننا نستخدم قانون المتوسط التوافقي للبيانات غير المبوبة، والذي يحسب كما يلي:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4}} = \frac{4}{\frac{1}{20} + \frac{1}{60} + \frac{1}{40} + \frac{1}{30}} = \frac{4}{0{,}125} = 32$$

إذن متوسط سرعة الدراج هي 32 كلم/سا

III–3–2–2–2. حساب المتوسط التوافقي للبيانات المبوبة:

إن المتوسط التوافقي كالمتوسط الهندسي عادة ما يستخدم عند ظواهر معينة لها نوع كمي مستمر، فتلك البيانات تكون عبارة عن معدلات سرعة أو متوسطات أسعار، أو متوسطات كثافة سكانية، وبالتالي يمكن القول أنه يستخدم في البيانات المبوبة للمتغير الكمي المستمر، ولا يمكن أن يحسب عند المتغير الكمي المتقطع لأنه قيم نقطية (صحيحة).

إذن إذا كانت مراكز الفئات هي : $C_1 \; ; \; C_2 \; ; \; .... \; ; \; C_k$ ، وكانت $n_1 \; ; \; n_2 \; ; \; .... \; ; \; n_k$ تمثل التكرارات المقابلة لها، فإن المتوسط التوافقي لهذه الفئات يحسب وفق القاعدة التالية:

$$H = \frac{n_1 + n_2 + ... + n_k}{\frac{n_1}{C_1} + \frac{n_2}{C_2} + ... + \frac{n_k}{C_k}}$$

$$H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{C_i}}$$

مثال (3–13): تمثل البيانات الموالية أسعار عشرون نوع من المواد الغذائية موزعة حسب مجالات سعرية.

الجدول رقم (3–14): توزيع المواد الغذائية حسب مجالات أسعارها (الوحدة : دج)

مجالات الأسعار	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$
عدد المواد الغذائية (n_i)	6	4	5	2	3

المطلوب: ما هو متوسط أسعار هذه المواد الغذائية ؟

الحل: بما أن البيانات متعلقة بمتوسطات الأسعار فالمتوسط الملائم هو المتوسط التوافقي.

الجدول رقم (3-15): حساب المتوسط التوافقي إعتمادا على جدول التوزيع التكراري

الفئات	n_i	C_i	$\frac{n_i}{C_i}$
$]20-10]$	6	15	0.4
$]30-20]$	4	25	0.16
$]40-30]$	5	35	0.14
$]50-40]$	2	45	0.04
$]60-50]$	3	55	0.05
المجموع	20	/	0.79

إذن متوسط أسعار المواد الغذائية هو :

$$H = \frac{\sum_{i=1}^{5} n_i}{\sum_{i=1}^{5} \frac{n_i}{C_i}} = \frac{20}{0.79} = 25.31$$

متوسط أسعار المواد الغذائية هو 25.31 دج

III-3-2-2-3. خواص المتوسط التوافقي:

يتميز المتوسط التوافقي بالخصائص التالية :

- نتيجة لصعوبة حساب المتوسط التوافقي فإنه أقل مقاييس النزعة المركزية إستخداما، ويقتصر أحيانا على إيجاد متوسطات الأسعار؛
- المتوسط التوافقي هو أقل المتوسطات $\overline{X} \geq G \geq H$ [1]
- عند حسابه يتم إستخدام جميع القيم ؛
- لا يمكن حسابه عند وجود بيانات صفرية (معدومة)، وكذلك من جداول التوزيع التكراري المفتوحة.

III-3-2-3. المتوسط التربيعي (Quadratic Mean):

يعتبر المتوسط التربيعي معقدا نوعا ما، فهو يستعمل في الحالات التي تتضمن بعض القيم السالبة والتي نريد معرفة متوسطها بغض النظر عن إشاراتها، ففي هذه الحالة لا يمكن إستخدام المتوسط الحسابي لأنه يلغي القيم السالبة مع الموجبة، ولا يمكن إستخدام الوسط الهندسي لأنه أيضا يوقعنا في البحث عن جذر سالب وهو أمر متعذر.[2]

[1] - محمد راتول ، مرجع سابق ، ص 126.

[2] - محمد بوهزة، مرجع سابق، ص 66.

يعرف المتوسط التربيعي لمجموعة من القيم على أنه الجذر التربيعي للمتوسط الحسابي لمربعات تلك القيم، ويرمز له بالرمز Q، وبما أنه يوجد نوعين من البيانات فأن الصيغة الرياضية له تختلف من بيانات إلى أخرى.

III-3-2-3-1. حساب المتوسط التربيعي للبيانات الأولية:

إذا كانت قيم الظاهرة المدروسة x_1 ; x_2 ; .... ; x_n والتي عددها **n** فإن متوسطها التربيعي يعطى بالصيغة التالية:

$$Q = \sqrt{\frac{{x_1}^2 + {x_2}^2 + ... + {x_n}^2}{n}}$$

$$Q = \sqrt{\frac{\sum_{i=1}^{n} {x_i}^2}{n}}$$

مثال (3-14): لتكن القيم الموالية : 20- ، 12 ، 24 ، 18 ، 22-

المطلوب: أحسب التوسط التربيعي للقيم ؟

الحل:

$$Q = \sqrt{\frac{{x_1}^2 + {x_2}^2 + ... + {x_n}^2}{n}}$$

لدينا:

$$Q = \sqrt{\frac{(-20)^2 + 12^2 + 24^2 + 18^2 + (-22)^2}{5}} = \sqrt{\frac{1928}{5}} = \sqrt{385{,}6} = 19{,}63$$

III-3-2-3-2. حساب المتوسط التربيعي للبيانات المبوبة:

بما أنه يوجد نوعين من المتغير الكمي، أحدهما متقطع (منفصل) والآخر مستمر (متصل) فإننا سنقوم بعرض الصيغة الرياضية للمتوسط التربيعي لكل نوع.

III-3-2-3-2-1. حساب المتوسط التربيعي لبيانات المتغير الكمي المتقطع:

إذا كانت القيم النقطية للظاهرة المدروسة هي : x_1 ; x_2 ; .... ; x_k ، وكانت n_1 ; n_2 ; .... ; n_k تمثل التكرارات المقابلة لها، فإن المتوسط التربيعي لهذه القيم يحسب وفق القاعدة التالية:

$$Q = \sqrt{\frac{(x_1)^2 n_1 + (x_2)^2 n_2 + ... + (x_k)^2 n_k}{n_1 + n_2 + ... + n_k}}$$

$$Q = \sqrt{\frac{\sum_{i=1}^{k} {x_i}^2 n_i}{\sum_{i=1}^{k} n_i}}$$

مثال (3-15): يحتوي جدول التوزيع التكراري الموالي على قيم متغير كمي متقطع (قيم نقطية)

الجدول رقم (3-15): جدول توزيع تكراري لبيانات متغير كمي متقطع

المتغير X_i	0	1	2	3	4
التكرار n_i	4	2	6	8	4

المطلوب: أحسب المتوسط التربيعي للبيانات ؟

الحل:

الجدول رقم (3-16): استخدام جدول التوزيع التكراري في حساب المتوسط التربيعي

المتغير X_i	**التكرار** n_i	x_i^2	$x_i^2 n_i$
0	4	0	0
1	2	1	2
2	6	4	24
3	8	9	72
4	4	16	64
$\sum$	24	-	162

إذن :

$$Q = \sqrt{\frac{(x_1)^2 n_1 + (x_2)^2 n_2 + ... + (x_k)^2 n_k}{n_1 + n_2 + ... + n_k}} = \sqrt{\frac{0^2 \times 4 + 1^2 \times 2 + 2^2 \times 6 + 3^2 \times 8 + 4^2 \times 4}{4+2+6+8+4}}$$

$$Q = \sqrt{\frac{162}{24}} = \sqrt{6,75} = 2.59 \approx 3$$

المتوسط التربيعي للقيم النقطية هو 3.

III-3-2-3-2-2. حساب المتوسط التربيعي لبيانات المتغير الكمي المستمر:

إذا كانت مراكز الفئات للمتغير المستمر هي : $C_1 ; C_2 ; .... ; C_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، فإن المتوسط التربيعي لهذا المتغير الكمي المستمر يحسب وفق القاعدة التالية:

$$Q = \sqrt{\frac{(C_1)^2 n_1 + (C_2)^2 n_2 + ... + (C_k)^2 n_k}{n_1 + n_2 + ... + n_k}}$$

$$Q = \sqrt{\frac{\sum_{i=1}^{k} C_i^2 n_i}{\sum_{i=1}^{k} n_i}}$$

مثال (3-17): يحتوي جدول التوزيع التكراري الموالي على فئات متغير كمي مستمر وتكراراتها المقابلة.

الجدول رقم (3-17): جدول توزيع تكراري لبيانات متغير كمي مستمر

الفئات	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$
التكرارات (n_i)	6	4	5	2	3

المطلوب: أحسب المتوسط التربيعي ؟

الحل:

الجدول رقم (3–18): حساب المتوسط التربيعي اعتمادا على جدول التوزيع التكراري

الفئات	n_i	C_i	C_i^2	$C_i^2 \times n_i$
$]20-10]$	6	15	225	1350
$]30-20]$	4	25	625	2500
$]40-30]$	5	35	1225	6125
$]50-40]$	2	45	2025	4050
$]60-50]$	3	55	3025	9075
المجموع	20	/	/	23100

إذن المتوسط التربيعي هو :

$$Q = \sqrt{\frac{(C_1)^2 n_1 + (C_2)^2 n_2 + ... + (C_k)^2 n_k}{n_1 + n_2 + ... + n_k}}$$

$$Q = \sqrt{\frac{23100}{20}} = \sqrt{1155} = 33.98$$

III–3–3. العلاقة بين المتوسطات الحسابية:

من خلال خواص المتوسطات التي ذكرناها نستنتج أن المتوسط التربيعي هو أكبر المتوسطات ثم يليه المتوسط الحسابي ثم المتوسط الهندسي، والأقل هو المتوسط التوافقي، أما إذا كانت القيم متساوية فإن المتوسطات تكون متساوية، أي:

$$H \leq G \leq \overline{X} \leq Q$$

مثال (3–18): بالرجوع إلى بيانات المثال رقم (3–01) الموالية :

2 ، 4 ، 8 ، 9 ، 10 ، 8 ، 10 ، 9 ، 10 ، 7 ، 5 ، 8

المطلوب : لقد وجدنا سابقا أن المتوسط الحسابي للقيم السابقة هو $\overline{X} = 7,5$ ، فأوجد كل من المتوسطات الأخرى، ثم قارن بينها ؟

الحل :

– حساب المتوسط التوافقي:

$$H=\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+....+\frac{1}{x_{12}}}=\frac{12}{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{8}+\frac{1}{10}+\frac{1}{9}+\frac{1}{10}+\frac{1}{7}+\frac{1}{5}+\frac{1}{8}}$$

$$H=\frac{12}{1.99}=6.03$$

- حساب المتوسط الهندسي:

$$G=\sqrt[n]{x_1\times x_2\times .....\times x_{12}}=\sqrt[12]{2\times 4\times 8\times 9\times 10\times 8\times 10\times 9\times 10\times 7\times 5\times 8}=\sqrt[12]{11612160000}$$

$$G=(11612160000)^{1/12}=6.89$$

- حساب المتوسط التربيعي:

$$Q=\sqrt{\frac{x_1^{\,2}+x_2^{\,2}+....+x_{12}^{\,2}}{n}}=\sqrt{\frac{2^2+4^2+8^2+9^2+10^2+8^2+10^2+9^2+10^2+7^2+5^2+8^2}{12}}$$

$$Q=\sqrt{\frac{748}{12}}=\sqrt{62.333}=7.89$$

إذن العلاقة محققة : $Q>\overline{X}>G>H$

مثال (3–19): لتكن القيم التالية 2 ، 2 ، 2 ، 2 ، 2

المطلوب : أحسب كل المتوسطات، ثم قارن بينها ؟

الحل :

- حساب المتوسط الحسابي:

$$\overline{X}=\frac{x_1+x_2+......+x_5}{n}=\frac{2+2+2+2+2}{5}=\frac{10}{5}=2$$

- حساب المتوسط التوافقي:

$$H=\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+....+\frac{1}{x_5}}=\frac{5}{\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+}=\frac{5}{\frac{5}{2}}=5\times\frac{2}{5}=2$$

- حساب المتوسط الهندسي:

$$G=\sqrt[n]{x_1\times x_2\times .....\times x_5}=\sqrt[5]{2\times 2\times 2\times 2\times 2}=\sqrt[5]{32}=(32)^{1/5}=2$$

- حساب المتوسط التربيعي:

$$Q=\sqrt{\frac{x_1^{\,2}+x_2^{\,2}+....+x_5^{\,2}}{n}}=\sqrt{\frac{2^2+2^2+2^2+2^2+2^2}{5}}=\sqrt{\frac{20}{5}}=\sqrt{4}=2$$

إذن المتوسطات متساوية وبالتالي العلاقة محققة $Q\geq\overline{X}\geq G\geq H$

III-3-4. الوسيط (The Median):

الوسيط هو تلك القيمة التي تقسم المجتمع الاحصائي إلى قسمين متساويين، بحيث تكون قيم المتغير الاحصائي مرتبة ترتيبا تصاعديا أو تنازليا [1] ، ويرمز له بالرمز M_e.

وبما أنه يوجد نوعين من البيانات، الأولية (غير المبوبة) والمبوبة فإن الوسيط يختلف في حسابه من نوع إلى آخر.

III-3-4-1. حساب الوسيط للبيانات الأولية (غير المبوبة):

لحساب الوسيط من البيانات الأولية يجب إتباع الخطوات التالية:

- أولا ترتيب البيانات (القيم) تصاعديا أو تنازليا؛
- ثانيا نقوم بمعرفة عدد البيانات (القيم)، لأن طريقة حساب الوسيط للبيانات غير المبوبة الفردية يختلف عنه للبيانات الزوجية.

III-3-4-1-1. حساب الوسيط للبيانات الأولية (غير المبوبة) المفردة:

عند حساب الوسيط للبيانات الأولية نرتب هذه البيانات ترتيبا تصاعديا أو تنازليا، وإذا كان عدد البيانات (القيم) (n) فردي فإن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$، أي $M_e = X_{\frac{n+1}{2}}$.

مثال (3-20): لتكن القيم التالية 5 ، 2 ، 7 ، 3 ، 6

المطلوب : أحسب الوسيط لهذه القيم؟

الحل :

نقوم أولا بترتيب القيم وليكن تنازليا

الجدول رقم (3-19): ترتيب القيم تنازليا وتحديد رتبها

الرتبة	X_1	X_2	X_3	X_4	X_5
ترتيب القيم تنازليا	**7**	**6**	**5**	**3**	**2**

لدينا عدد القيم فردي وبالتالي رتبة الوسيط هي : $\frac{n+1}{2} = \frac{5+1}{2} = \frac{6}{2} = 3$ وبالتالي الوسيط هو القيمة التي ترتيبها يقابل رتبة الوسيط، ويتم إستخراجه كما يلي: $M_e = X_{\frac{n+1}{2}} = X_{\frac{5+1}{2}} = X_3 = 5$.

III-3-4-1-2. حساب الوسيط للبيانات الأولية (غير المبوبة) الزوجية:

إذا كان عدد القيم (n) زوجي فإننا نصادف قيمتين في وسط البيانات المرتبة تصاعديا أو تنازليا، فرتبتي هاتين القيمتين هما : ($\frac{n}{2}, \frac{n}{2}+1$)، وبالتالي قيمة الوسيط تساوي المتوسط الحسابي لهاتين القيمتين، أي: ($M_e = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$).

[1] - جيلالي جلاطو، مرجع سابق، ص 41.

مثال (3-21): لتكن القيم التالية 20 ، 22 ، 21 ، 27 ، 25 ، 26

المطلوب : أحسب الوسيط لهذه القيم؟

الحل :

نقوم أولا بترتيب القيم وليكن تصاعديا

الجدول رقم (3-20): ترتيب القيم تصاعديا وتحديد رتبها

الرتبة	X_1	X_2	X_3	X_4	X_5	X_6
ترتيب القيم تصاعديا	20	21	22	24	26	27

لدينا عدد القيم زوجي، وبالتالي رتبتي القيمتين الوسطيتين هما : $(\frac{6}{2},\frac{6}{2}+1)=(3;4)$ ، والوسيط هو المتوسط الحسابي للقيمتين

اللتين رتبتيهما السابقتين $(3;4)$ ، ويتم إستخراجه كما يلي:

$$M_e=\frac{X_{\frac{n}{2}}+X_{\frac{n}{2}+1}}{2}=\frac{X_{\frac{6}{2}}+X_{\frac{6}{2}+1}}{2}=\frac{X_3+X_4}{2}=\frac{22+24}{2}=23$$

III-3-4-2. حساب الوسيط للبيانات المبوبة:

عادة ما نصادف نوعين من البيانات الكمية المبوبة، بيانات نقطية (متغير كمي متقطع) وبيانات ذات فئات (متغير كمي مستمر)، والوسيط يختلف في حسابه من نوع إلى آخر.

III-3-4-2-1. حساب الوسيط لبيانات المتغير الكمي المنفصل (المتقطع):

من أجل حساب الوسيط لبيانات مبوبة في حالة متغير كمي منفصل، نتبع الخطوات التالية:

1- نحساب التكرار المتجمع الصاعد؛

2- نحدد رتبة الوسيط وفق القاعدة $(\frac{N}{2})$، حيث $N=\sum_{i=1}^{k}n_i$ ؛

3- نحدد في العمود الخاص بالتكرار المتجمع الصاعد من الأعلى قيمة تساوي رتبة الوسيط أو أعلى منها مباشرة، وهذه القيمة تقابلها في المتغيرة قيمة الوسيط.

مثال (3-21): لتكن البيانات الموالية والتي تمثل معطيات المثال رقم (3-04).

الجدول رقم (3-21): توزيع مجموعة من عمال مؤسسة حسب عدد أيام التغيب.

عدد أيام الغياب X_i	0	1	2	3	4	5	$\sum$
عدد العمال n_i	10	8	4	2	3	3	30

المطلوب: أحسب الوسيط ؟

الحل:

لحساب الوسيط لهذه البيانات نتبع الخطوات السابقة وهي حساب التكرار التجميعي الصاعد والبحث فيه عن مكان رتبة الوسيط.

الجدول رقم (3-21): استخدام جدول التوزيع التكراري في حساب الوسيط

عدد أيام التغيب X_i	عدد العمال n_i	$n_i \uparrow$
0	10	10
1	8	18
2	4	22
3	2	24
4	3	27
5	3	30
$\sum$	30	-

- نحسب رتبة الوسيط وهي : $\frac{N}{2}=\frac{30}{2}=15$ ، ثم نبحث في قيم العمود الخاص بالتكرار المتجمع الصاعد من القيمة الأولى (10) وهي قيمة أقل من قيمة الرتبة ولا تحقق الشرط، ثم القيمة اللاحقة (18) وهي قيمة أكبر من رتبة الوسيط، وهي التي تحقق الشرط (لابد أن تكون القيمة تساوي رتبة الوسيط أو أكبر منها مباشرة)، وبالتالي القيمة المقابلة في المتغيرة هي قيمة الوسيط، أي أن: $M_e=1$

III-3-4-2-2-2. حساب الوسيط لبيانات المتغير الكمي المتصل (المستمر):

من أجل تحديد الوسيط لبيانات مبوبة لمتغير كمي متصل، نتبع الخطوات التالية:

1- نحسب التكرار المتجمع الصاعد؛

2- نحسب رتبة الوسيط وفق العلاقة $(\frac{N}{2})$، حيث $N=\sum_{i=1}^{k} n_i$ ؛

3- نستخرج الفئة الوسيطية وهي الفئة التي ينتمي إليها الوسيط، وهي الفئة المقابلة للخانة التي تتواجد فيها رتبة الوسيط في التكرار المتجمع الصاعد كما حددنا ذلك سابقا.

4- نحسب الوسيط بتطبيق العلاقة الرياضية التالية:

$$M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_e} \times k$$

حيث : L_1 هو الحد الأدنى للفئة الوسيطية

N يمثل مجموع التكرارات $N=\sum_{i=1}^{k} n_i$

N_0 هو التكرار المتجمع الصاعد للفئة قبل الوسيطية

n_e هو تكرار الفئة الوسيطية

k يمثل طول الفئة الوسيطية

مثال (3–22): لتكن البيانات الإحصائية الموالية التي تمثل معطيات المثال السابق رقم (**3–06**)

الجدول رقم (3–22): توزيع مجموعة من الأفراد حسب دخلهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: أحسب الوسيط لهذه البيانات ؟

الحل: إعتمادا على جدول التوزيع التكراري يتم حساب الوسيط وهذا بعد حساب التكرار التجميعي الصاعد.

الجدول رقم (3–23): حساب الوسيط اعتمادا على جدول التوزيع التكراري.

فئات الدخل الشهري X_i	عدد الأفراد n_i	التكرار التجميعي الصاعد $n_i \uparrow$
$]20-10]$	5	5
$]30-20]$	8	13
$]40-30]$	12	25
$]50-40]$	25	50
$]60-50]$	10	60
المجموع	60	/

بعد حساب التكرار التجميعي الصاعد المبين في الجدول نقوم باتباع الخطوات التالية:

- نحسب رتبة الوسيط وفق القاعدة : $\frac{N}{2}=\frac{60}{2}=30$ ، ثم نبحث في قيم التكرار التجميعي الصاعد بداية من القيمة الأولى عن قيمة تساوي هذه الرتبة أو أكبر منها مباشرة، وبالتالي القيم الأولى (5 و 13 و 25) أقل من الرتبة ولا يمكن أن نختار أي منهم، ونلاحظ أن رتبة الوسيط تدخل ضمن القيمة (50) الموالية، وبالتالي الفئة المقابلة لذلك هي الفئة الوسيطية $]50-40]$.
- بعد تحديد الفئة الوسيطية نحسب الوسيط وفق القاعدة السابقة :

$$M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_e} \times k = 40 + \frac{\frac{60}{2} - 25}{25} \times 10 = 40 + \frac{50}{25} = 42$$

$$M_e = 42 \times 10^3 \quad DA$$

III-3-4-3. استخراج الوسيط بيانيا:

إذا كانت لدينا بيانات مبوبة لمتغير كمي مستمر، فيمكن استخراج وسيطها بيانيا وهذا من خلال منحنى التكرار المتجمع الصاعد والنازل معا ، أو من خلال إحداهما.

- إذا كان مثلا لدينا أحد المنحنيين الصاعد أو النازل:
  - نرسم التكرار التجمعي الصاعد أو النازل كما وضحنا ذلك في المحور السابق.
  - نحدد رتبة الوسيط كنقطة في المحور العمودي، أي محور التراتيب (محور التكرار التجميعي الصاعد أو النازل)، ثم نسقط هذه النقطة أفقيا على منحنى التكرار التجميعي الصاعد أو النازل ، وعند نقطة التماس نعيد إسقاط آخر عمودي على محور المتغير، ونقطة التقاطع الأخيرة بين محور الاسقاط ومحور المتغير هي قيمة الوسيط.
- أما إذا كان لدينا رسم بياني يحتوي على منحنيي التكرار التجميعي الصاعد والنازل، فإننا نسقط نقط تقاطعهما على محور المتغير لنحصل على قيمة الوسيط

مثال (3-23): لتكن بيانات الجدول رقم (**3-23**) السابق

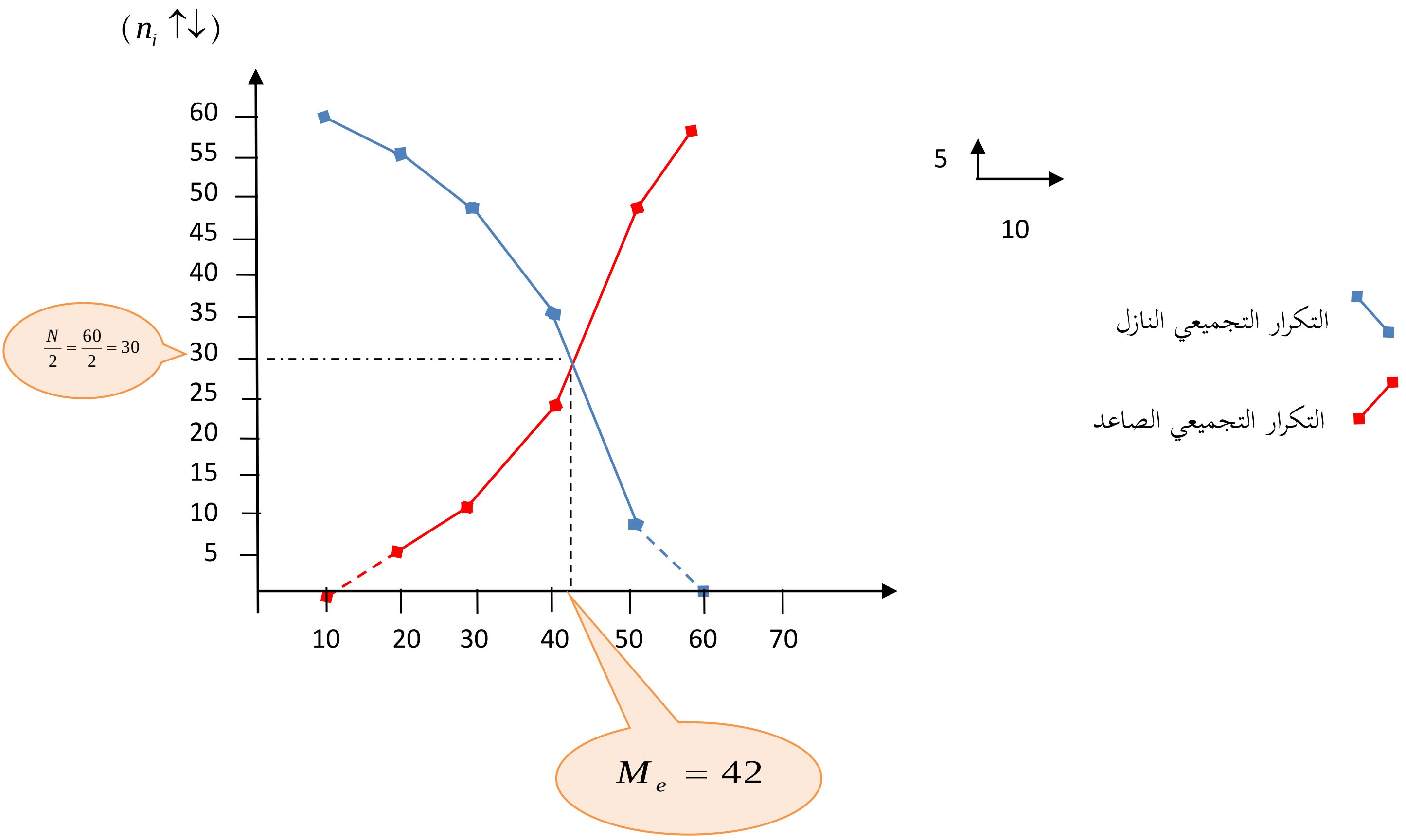
المطلوب : أحسب التكرار التجميعي النازل، ثم استخدم معطيات الجدول في استخراج الوسيط بيانيا ؟

الحل:

الجدول رقم (3-24): حساب التكرارين التجميعيين الصاعد $n_i \uparrow$ والنازل $n_i \downarrow$

فئات الدخل الشهري X_i	عدد الأفراد n_i	التكرار التجميعي الصاعد $n_i \uparrow$	التكرار التجميعي النازل $n_i \downarrow$
$]20-10]$	5	5	60
$]30-20]$	8	13	55
$]40-30]$	12	25	47
$]50-40]$	25	50	35
$]60-50]$	10	60	10
المجموع	60	/	/

لقد قمنا بحساب رتبة الوسيط سابقا ووجدناها $\frac{N}{2}=\frac{60}{2}=30$ ، وقمنا بحساب الوسيط فوجدناه $M_e = 42$



III-3-4-4. خواص الوسيط:

للوسيط خواص نذكرها فيما يلي:

- يتغير الوسيط كلما غيرنا أطوال الفئات بالنسبة لنفس التوزيع التكراري، وبالتالي الوسيط يتميز بعدم الثبات؛
- لا يتأثر الوسيط بالقيم المتطرفة أو الشاذة؛ [1]

وبناء على الأمثلة التي قمنا بها، يمكن القول أن الوسيط :

- لا يستخدم كل البيانات؛
- يمكن حسابه من جداول التوزيع التكرارية المفتوحة؛
- يمكن استخراج قيمته بيانيا؛

III-3-5. أشباه الوسيط (الربيعيات، العشيريات والمئويات)

من خلال تعريف الوسيط قلنا أنه القيمة التي تقسم مجموع البيانات إلى قسمين متساويين، كما يمكن تقسيم هذه البيانات إلى عدة أقسام متساوية، قد تكون أربعة والتي تسمى بالربيعيات، أو عشرة والتي تسمى بالعشيرات، أو مائة والتي تسمى بالمئويات.

[1] - جلاطو جيلالي، مرجع سابق، ص 45.

الربيعيات (Quartile): هي القيم التي تقسم البيانات المرتبة تصاعديا أو تنازليا إلى أربع أجزاء متساوية، حيث كل قسم يمثل 25 % من البيانات، وعدد هذه الربيعيات (Q_i) هو ثلاثة، حيث i=1 ;2 ;3 ، فالربيع الأول يقسم البيانات إلى 25 % قبله و 75 % بعده، أما الربيع الثاني فيسبقه 50 % من عدد البيانات ونفس النسبة بعده، أما الربيع الثالث فيسبقه 75 % من عدد البيانات ويليه 25% منها.

العشيريات (Decile): هي القيم التي تقسم البيانات المرتبة تصاعديا أو تنازليا إلى عشرة أجزاء متساوية، حيث كل قسم يمثل 10 % من البيانات، وعدد هذه العشيريات (D_i) هو تسعة، حيث i=1 ;2 ;3 ;.... ;9 ، فالعشير الأول يقسم البيانات إلى 10 % قبله و 90 % بعده، أما العشير الثاني فيسبقه 20 % من البيانات ويليه 80% من عدد البيانات،....وهكذا حتى العشير التاسع، الذي يسبقه 90% من عدد البيانات وبعده 10% منها.

المئويات (Percentile): هي القيم التي تقسم البيانات المرتبة تصاعديا أو تنازليا إلى مائة جزء متساوي، حيث كل جزء يمثل 1 % من البيانات، وعدد هذه المئويات (P_i) هو مائة، حيث i=1 ;2 ;3 ;.... ;99 ، فالمئوي الأول يقسم البيانات إلى 1 % قبله و 99 % بعده، أما العشير الثاني فيسبقه 2 % من البيانات ويليه 98% من عدد البيانات،....وهكذا حتى المئوي التاسع والتسعون، الذي يسبقه 99% من عدد البيانات وبعده 1% منها.

ملاحظة: تتساوى بعض الربيعيات مع العشيريات والمئويات، فالربيع الثاني يتساوى مع الوسيط ومع العشير الخامس ومع المئوي الخمسون، كما يتساوى كذلك الربيع الأول مع المئوي الخامس والعشرون، كما يتساوى كذلك الربيع الثالث مع المئوي الخامس والسبعون، وتتساوى كل العشيريات مع المئويات مقابلة لها، مثل العشير الأول مع المئوي العاشر، والعشير الثاني مع المئوي العشرون.........، والعشير التاسع مع المئوي التسعون.

III-3-5-1. الربيعيات (Quartile):

III-3-5-1-1. حساب الربيعيات للبيانات الأولية (غير المبوبة):

إن منهجية حساب الربيعيات تتشابه مع طريقة حساب الوسيط للبيانات غير المبوبة، حيث :

- يتم ترتيب البيانات ترتيبا تصاعديا.
- فإذا كان عدد البيانات فردي فقيمة الربيعيات هي القيم ذات الرتب $\frac{i(n+1)}{4}$ ، حيث i =1 ;2 ;3 وهي تمثل رقم الربيع، وقيمة الربيع (Q_i) هي : $Q_i = X_{\frac{i(n+1)}{4}}$
- أما إذا كان عدد البيانات زوجي فإن الربيعيات هي المتوسط الحسابي للقيمتين اللتين ترتيبهما ($\frac{in}{4}$) و ($\frac{in}{4}+1$)، أي:

$$Q_i = \frac{X_{\frac{in}{4}} + X_{\frac{in}{4}+1}}{2}$$

- حتى يتم إستخراج الربيعيات بصورة واضحة وسليمة في البيانات غير المبوبة من الأفضل أن يكون عدد البيانات كالتالي :

- في حالة عدد البيانات فردي: تؤخذ عدد البيانات كحدود متتالية حسابية حدها الأول 7 وأساسها 4 ، أي عدد القيم الفردية يكون وفق التسلسل التالي : 7 ، 11 ، 15 ، 19.............
- في حالة عدد البيانات زوجي: تؤخذ عدد البيانات كحدود متتالية حسابية حدها الأول 8 وأساسها 4 ، أي عدد القيم الزوجية يكون وفق التسلسل التالي : 8 ، 12 ، 16 ، 20.............

مثال (3–24): تمثل البيانات الموالية سلسلتين من الأعداد

السلسلة الأولى : 2 ، 1 ، 5 ، 6 ، 3 ، 2 ، 4 ، 7 ، 8 ، 9 ، 8

السلسلة الثانية : 10 ، 11 ، 13 ، 19 ، 21 ، 14 ، 15 ، 17 ، 12 ، 15 ، 19 ، 16

المطلوب: حدد ربيعات البيانات في السلسلتين ؟

الحل:

السلسلة الأولى : 2 ، 1 ، 5 ، 5 ، 6 ، 1 ، 3 ، 2 ، 4 ، 7 ، 8 ، 9 ، 9 ، 7 ، 8

نلاحظ أن عدد بيانات السلسلة الأولى فردي (11)، وقبل حساب الربيعات لابد من ترتيب القيم تصاعديا

الجدول رقم (3–25): ترتيب القيم تصاعديا وتحديد رتبها

الرتبة	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
ترتيب القيم تصاعديا	**1**	**2**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**8**	**9**

- حساب الربيع الأول (Q_1):

بما أن عدد البيانات فردي، فرتبة الربيع الأول هي $\frac{i(n+1)}{4}=\frac{1(11+1)}{4}=\frac{12}{4}=3$

إذن قيمة الربيع الأول هي : $Q_1=X_{\frac{i(n+1)}{4}}=X_{\frac{1(11+1)}{4}}=X_{\frac{12}{4}}=X_3=2$

- حساب الربيع الثاني (Q_2):

رتبة الربيع الثاني هي: $\frac{i(n+1)}{4}=\frac{2(11+1)}{4}=\frac{24}{4}=6$

إذن قيمة الربيع الثاني هي : $Q_2=X_{\frac{i(n+1)}{4}}=X_{\frac{2(11+1)}{4}}=X_{\frac{24}{4}}=X_6=5$

- حساب الربيع الثالث (Q_3):

رتبة الربيع الثالث هي: $\frac{i(n+1)}{4}=\frac{3(11+1)}{4}=\frac{36}{4}=9$

إذن قيمة الربيع الثالث هي : $Q_2=X_{\frac{i(n+1)}{4}}=X_{\frac{3(11+1)}{4}}=X_{\frac{36}{4}}=X_9=8$

السلسلة الثانية : 10 ، 11 ، 13 ، 19 ، 21 ، 14 ، 15 ، 17 ، 12 ، 15 ، 19 ، 16

نلاحظ أن عدد بيانات السلسلة الأولى زوجي (12)، وقبل حساب الربيعيات لابد من ترتيب القيم تصاعديا.

الجدول رقم (3-25): ترتيب القيم تصاعديا وتحديد رتبها

الرتبة	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
ترتيب القيم تصاعديا	**10**	**11**	**12**	**13**	**14**	**15**	**15**	**16**	**17**	**19**	**19**	**21**

- حساب الربيع الأول (Q_1):

بما أن عدد البيانات زوجي، فرتبتي الربيع الأول هما : $\frac{in}{4}+1=\frac{1(12)}{4}+1=4 \quad et \quad \frac{in}{4}=\frac{1(12)}{4}=\frac{12}{4}=3$

إذن قيمة الربيع الأول هي : $Q_i=\frac{X_{\frac{in}{4}}+X_{\frac{in}{4}+1}}{2}=\frac{X_{\frac{1(12)}{4}}+X_{\frac{1(12)}{4}+1}}{2}=\frac{X_3+X_4}{2}=\frac{12+13}{2}=12.5$

- حساب الربيع الثاني (Q_2):

رتبتي الربيع الثاني هما: $\frac{in}{4}+1=\frac{2(12)}{4}+1=7 \quad et \quad \frac{in}{4}=\frac{2(12)}{4}=\frac{24}{4}=6$

إذن قيمة الربيع الثاني هي : $Q_i=\frac{X_{\frac{in}{4}}+X_{\frac{in}{4}+1}}{2}=\frac{X_{\frac{2(12)}{4}}+X_{\frac{2(12)}{4}+1}}{2}=\frac{X_6+X_7}{2}=\frac{15+15}{2}=15$

- حساب الربيع الثالث (Q_3):

رتبتي الربيع الثالث هما: $\frac{in}{4}+1=\frac{3(12)}{4}+1=10 \quad et \quad \frac{in}{4}=\frac{3(12)}{4}=\frac{36}{4}=9$

إذن قيمة الربيع الثالث هي : $Q_i=\frac{X_{\frac{in}{4}}+X_{\frac{in}{4}+1}}{2}=\frac{X_{\frac{3(12)}{4}}+X_{\frac{3(12)}{4}+1}}{2}=\frac{X_9+X_{10}}{2}=\frac{17+19}{2}=18$

III-3-5-1-2. حساب الربيعيات للبيانات المبوبة:

بما أنه يوجد نوعين من البيانات الكمية المبوبة، بيانات متعلقة بالمتغير الكمي المنقطع وأخرى بالمستمر فإننا سنقوم بطريقة إستخراج وحساب الربيعيات حسب نوع من هذه البيانات.

III-3-5-1-2-1. حساب الربيعيات لبيانات المتغير الكمي المنقطع:

إن إستخراج الربيعيات من بيانات المتغير الكمي المتقطع يتبع نفس الخطوات المتبعة في إستخراج الوسيط، حيث هذه الخطوات هي:

1- نحسب التكرارات التجميعية الصاعدة؛

2- نحدد رتبة الربيع i المعطاة بالعلاقة التالية: $\frac{i\sum n_i}{4}=\frac{iN}{4}$؛

3- نحدد في العمود الخاص بالتكرار المتجمع الصاعد من الأعلى قيمة تساوي رتبة الربيع i أو أعلى منها مباشرة، وهذه القيمة تقابلها في المتغيرة قيمة الربيع Q_i .

مثال (3-25): لتكن البيانات الموالية والتي تمثل معطيات المثال رقم (3-21).

الجدول رقم (3-26): توزيع مجموعة من عمال مؤسسة حسب عدد أيام التغيب.

عدد أيام الغياب X_i	0	1	2	3	4	5	$\sum$
عدد العمال n_i	10	8	4	2	3	3	30

المطلوب: أحسب الربيعيات ؟

الحل:

لحساب الربيعيات لهذه البيانات نحسب التكرار التجميعي الصاعد ثم نبحث فيه عن مكان رتبة كل ربيع.

الجدول رقم (3-27): استخدام جدول التوزيع التكراري في حساب الربيعيات

عدد أيام التغيب X_i	**عدد العمال n_i**	$n_i\uparrow$
0	10	10
1	8	18
2	4	22
3	2	24
4	3	27
5	3	30
$\sum$	30	-

- حساب الربيع الأول (Q_1) :

نحسب رتبة الربيع الأول وهي : $\frac{iN}{4}=\frac{1(30)}{4}=7,5$ ، ثم نبحث في قيم العمود الخاص بالتكرار المتجمع الصاعد من القيمة الأولى (10) وهي قيمة أكبر من قيمة الرتبة وهي التي تحقق الشرط (لابد أن تكون القيمة تساوي رتبة الربيع أو أكبر منها مباشرة)، وبالتالي القيمة المقابلة في المتغيرة هي قيمة الربيع الأول، أي أن: $Q_1=0$

- حساب الربيع الثاني (Q_2) :

نحسب رتبة الربيع الثاني وهي : $\frac{iN}{4}=\frac{2(30)}{4}=15$ ، ثم نبحث في قيم العمود الخاص بالتكرار المتجمع الصاعد من القيمة الأولى (10) وهي قيمة أصغر من رتبة الربيع الثاني، وبالتالي هي قيمة مرفوضة، ثم ننتقل إلى القيمة الأسفل منها (18)، وهي قيمة أكبر من قيمة الرتبة (15) وهي التي تحقق الشرط (لابد أن تكون القيمة تساوي رتبة الربيع أو أكبر منها مباشرة)، وبالتالي القيمة المقابلة في المتغيرة هي قيمة الربيع الثاني، أي أن: $Q_2=1$

- حساب الربيع الثالث (Q_3) :

نحسب رتبة الربيع الثالث وهي : $\frac{iN}{4} = \frac{3(30)}{4} = 22{,}5$ ، ثم نبحث في قيم العمود الخاص بالتكرار المتجمع الصاعد من القيم الأولى (10 ، 18 ، 22) وهي قيم أصغر من رتبة الربيع الثالث، وبالتالي هي قيم مرفوضة، ثم ننتقل إلى القيمة الأسفل منهم (24)، وهي قيمة أكبر من قيمة الرتبة (22,5) وهي التي تحقق الشرط (لابد أن تكون القيمة تساوي رتبة الربيع أو أكبر منها مباشرة)، وبالتالي القيمة المقابلة في المتغيرة هي قيمة الربيع الثالث، أي أن: $Q_3 = 3$:

III-3-5-1-2-2. حساب الربيعيات لبيانات المتغير الكمي المستمر (المتصل):

من أجل حساب الربيعيات (Q_i)، حيث $i = 1 ;2 ;3$ نتبع نفس خطوات حساب الوسيط وهي كالتالي:

1- نحسب التكرارات التجميعية الصاعدة؛

2- نحدد رتبة الربيع i المعطاة بالعلاقة التالية: $\frac{iN}{4} = \frac{i\sum n_i}{4}$ ؛

3- نحدد في العمود الخاص بالتكرار المتجمع الصاعد من الأعلى قيمة تساوي رتبة الربيع i أو أعلى منها مباشرة، وهذه القيمة تقابلها في المتغيرة فئة الربيع Q_i .

4- نحسب قيمة الربيع وفق القاعدة التالية:

$$Q_i = L_1 + \frac{\frac{iN}{4} - N_0}{n_{Q_i}} \times k$$

حيث : L_1 هو الحد الأدنى للفئة الربيعية؛

N يمثل مجموع التكرارات $N = \sum_{i=1}^{k} n_i$

N_0 هو التكرار المتجمع الصاعد للفئة قبل الربيعية؛

n_{Qi} هو تكرار الفئة الربيعية

k يمثل طول الفئة الربيعية

مثال (3-26): لتكن البيانات الإحصائية الموالية التي تمثل معطيات المثال السابق رقم **(3-06)**

الجدول رقم (3-28): توزيع مجموعة من الأفراد حسب دخلهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	$]10-20]$	$]20-30]$	$]30-40]$	$]40-50]$	$]50-60]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: أحسب الربيعيات الثلاثة لهذه البيانات ؟

الحل: باستخدام جدول التوزيع التكراري يتم حساب الربيعيات (Q_i) وهذا بعد حساب التكرار التجميعي الصاعد.

الجدول رقم (3–29): حساب التكرار التجميعي الصاعد اعتمادا على جدول التوزيع التكراري.

فئات الدخل الشهري X_i	عدد الأفراد n_i	التكرار التجميعي الصاعد $n_i \uparrow$
$]20-10]$	5	5
$]30-20]$	8	13
$]40-30]$	12	25
$]50-40]$	25	50
$]60-50]$	10	60
المجموع	60	/

– حساب الربيع الأول (Q_1):

➢ نحسب رتبة الربيع الأول (Q_1) وفق القاعدة : $\frac{iN}{4}=\frac{1(60)}{4}=15$، ثم نبحث في قيم التكرار التجميعي الصاعد بداية من القيمة الأولى عن قيمة تساوي هذه الرتبة أو أكبر منها مباشرة، وبالتالي القيم الأولى (5 و 13) أقل من الرتبة ولا يمكن أن نختار أي منهم، ونلاحظ أن رتبة الربيع تدخل ضمن القيمة (25) الموالية، وبالتالي الفئة المقابلة لذلك هي فئة الربيع الأول $]40-30]$

➢ بعد تحديد فئة الربيع الأول نحسب قيمة هذا الأخير وفق القاعدة السابقة:

$$Q_1 = L_1 + \frac{\frac{iN}{4} - N_0}{n_{Q_1}} \times k = 30 + \frac{\frac{1(60)}{4} - 13}{12} \times 10 = 30 + \frac{15-13}{12} \times 10 = 31{,}66$$

– حساب الربيع الثاني (Q_2):

➢ نحسب رتبة الربيع الثاني (Q_2) وفق القاعدة : $\frac{iN}{4}=\frac{2(60)}{4}=30$، ثم نبحث في قيم التكرار التجميعي الصاعد بداية من القيمة الأولى عن قيمة تساوي هذه الرتبة أو أكبر منها مباشرة، وبالتالي القيم الأولى (5 و 13 و 25) أقل من الرتبة ولا يمكن أن نختار أي منهم، ونلاحظ أن رتبة الربيع الثاني تدخل ضمن القيمة (50) الموالية، وبالتالي الفئة المقابلة لذلك هي فئة الربيع الثاني $]50-40]$

➢ بعد تحديد فئة الربيع الثاني نحسب قيمة هذا الأخير وفق القاعدة السابقة:

$$Q_2 = L_1 + \frac{\frac{iN}{4} - N_0}{n_{Q_2}} \times k = 40 + \frac{\frac{2(60)}{4} - 25}{25} \times 10 = 40 + \frac{30-25}{25} \times 10 = 42$$

– حساب الربيع الثالث (Q_3):

➢ نحسب رتبة الربيع الثالث (Q_3) وفق القاعدة : $\frac{iN}{4}=\frac{3(60)}{4}=45$، ثم نبحث في قيم التكرار التجميعي الصاعد بداية من القيمة الأولى عن قيمة تساوي هذه الرتبة أو أكبر منها مباشرة، وبالتالي القيم الأولى (5 و 13 و 25) أقل من الرتبة ولا يمكن

أن نختار أي منهم، ونلاحظ أن رتبة الربيع الثالث تدخل ضمن القيمة (50) الموالية، وبالتالي الفئة المقابلة لذلك هي فئة الربيع الثالث $[40-50[$ وهي نفس فئة الربيع الثاني.

➢ بعد تحديد فئة الربيع الثالث نحسب قيمة هذا الأخير وفق القاعدة السابقة:

$$Q_3 = L_1 + \frac{\frac{iN}{4} - N_0}{n_{Q_3}} \times k = 40 + \frac{\frac{3(60)}{4} - 25}{25} \times 10 = 40 + \frac{45-25}{25} \times 10 = 48$$

بالرغم من أن للربيعين الثاني والثالث نفس الفئة الربيعية إلا أن قيمتيهما مختلفتين وهذا راجع إلى أن لكل منهما رتبة مختلفة عن الأخرى.

III-3-5-2. العشيريات (Decile) والمئويات (Percentile):

هناك تشابه كبير ما بين العشيريات والمئويات وبالتالي إرتأينا أن ندرجهما معا.

III-3-5-2-1. حساب العشيريات والمئويات للبيانات الأولية (غير المبوبة):

إن منهجية حساب العشيريات والمئويات تتشابه مع طريقة حساب الربيعيات للبيانات غير المبوبة، حيث :

- يتم ترتيب البيانات ترتيبا تصاعديا.
- إذا كان عدد البيانات فردي فقيمة العشيريات هي القيم ذات الرتب $\frac{i(n+1)}{10}$ ، حيث i =1; 2;.... 9; وهي تمثل رقم العشير، وقيمة العشير (D_i) هي : $D_i = X_{\frac{i(n+1)}{10}}$ ، أما المئويات فهي القيم ذات الرتب $\frac{i(n+1)}{100}$ ، حيث i =1; 2;.... ;99 وهي تمثل رقم المئوي، وقيمة المئوي (P_i) هي : $P_i = X_{\frac{i(n+1)}{100}}$
- إذا كان عدد البيانات زوجي فإن العشيريات هي المتوسط الحسابي للقيمتين اللتين ترتيبهما $(\frac{in}{10})$ و $(\frac{in}{10}+1)$، أي: $D_i = \frac{X_{\frac{in}{10}} + X_{\frac{in}{10}+1}}{2}$ ، أما المئويات هي المتوسط الحسابي للقيمتين اللتين ترتيبهما $(\frac{in}{100})$ و $(\frac{in}{100}+1)$، أي:
$$P_i = \frac{X_{\frac{in}{100}} + X_{\frac{in}{100}+1}}{2}$$
- حتى يتم إستخراج العشيريات بصورة واضحة ولا تكون متداخلة فيما بينها في البيانات غير المبوبة من الأفضل أن يكون عدد البيانات كالتالي:

➢ في حالة عدد البيانات فردي: تؤخذ عدد البيانات كحدود متتالية حسابية حدها الأول 19 وأساسها 10 ، أي عدد القيم الفردية يكون وفق التسلسل التالي : 19 ، 29 ، 39 ، 49.............

➢ في حالة عدد البيانات زوجي: تؤخذ عدد البيانات كحدود متتالية حسابية حدها الأول 20 وأساسها 10 ، أي عدد القيم الزوجية يكون وفق التسلسل التالي : 20 ، 30 ، 40 ، 50.............

– وحتى يتم إستخراج المئويات بصورة واضحة ولا تكون متداخلة كذلك فيما بينها عند البيانات غير المبوبة من الأفضل أن يكون عدد البيانات كالتالي:

➢ في حالة عدد البيانات فردي: تؤخذ عدد البيانات كحدود متتالية حسابية حدها الأول 199 وأساسها 100 ، أي عدد القيم الفردية يكون وفق التسلسل التالي : 199 ، 299 ، 399 ، 499.............

➢ في حالة عدد البيانات زوجي: تؤخذ عدد البيانات كحدود متتالية حسابية حدها الأول 200 وأساسها 100 ، أي عدد القيم الزوجية يكون وفق التسلسل التالي : 200 ، 300 ، 400 ، 500.............

مثال (3–27): لتكن سلسلة الأعداد الطبيعية من 1 إلى 199

المطلوب: – أكتب سلسلة الأعداد مرتبة تصاعديا ثم حدد العشيريات والمئويات، وماذا تلاحظ ؟

– بإستخدام القواعد الرياضية السابقة أحسب العشير الخامس والمئوي السبعون، ثم تأكد من ذلك وفق ما توصلت إليه سابقا.

الحل:

➢ كتابة السلسلة وتحديد العشيريات والمئويات : نحدد العشيريات بـ .. والمئويات بـ ..

1 ، 2 ، 3 ، 4 ، 5 ، 6 ، 7 ، 8 ، 9 ، 10 ، 11 ، 12 ، 13 ، 14 ، 15 ، 16 ، 17 ، 18 ، 19 ، 20 ، 21 ، 22 ، 23 ، 24 ، 25 ، 26 ، 27 ، 28 ، 29 ، 30 ، 31 ، 32 ، 33 ، 34 ، 35 ، 36 ، 37 ، 38 ، 39 ، 40 ، 41 ، 42 ، 43 ، 44 ، 45 ، 46 ، 47 ، 48 ، 49 ، 50 ، 51 ، 52 ، 53 ، 54 ، 55 ، 56 ، 57 ، 58 ، 59 ، 60 ، 61 ، 62 ، 63 ، 64 ، 65 ، 66 ، 67 ، 68 ، 69 ، 70 ، 71 ، 72 ، 73 ، 74 ، 75 ، 76 ، 77 ، 78 ، 79 ، 80 ، 81 ، 82 ، 83 ، 84 ، 85 ، 86 ، 87 ، 88 ، 89 ، 90 ، 91 ، 92 ، 93 ، 94 ، 95 ، 96 ، 97 ، 98 ، 99 ، 100 ، 101 ، 102 ، 103 ، 104 ، 105 ، 106 ، 107 ، 108 ، 109 ، 110 ، 111 ، 112 ، 113 ، 114 ، 115 ، 116 ، 117 ، 118 ، 119 ، 120 ، 121 ، 122 ، 123 ، 124 ، 125 ، 126 ، 127 ، 128 ، 129 ، 130 ، 131 ، 132 ، 133 ، 134 ، 135 ، 136 ، 137 ، 138 ، 139 ، 140 ، 141 ، 142 ، 143 ، 144 ، 145 ، 146 ، 147 ، 148 ، 149 ، 150 ، 151 ، 152 ، 153 ، 154 ، 155 ، 156 ، 157 ، 158 ، 159 ، 160 ، 161 ، 162 ، 163 ، 164 ، 165 ، 166 ، 167 ، 168 ، 169 ، 170 ، 171 ، 172 ، 173 ، 174 ، 175 ، 176 ، 177 ، 178 ، 179 ، 180 ، 181 ، 182 ، 183 ، 184 ، 185 ، 186 ، 187 ، 188 ، 189 ، 190 ، 191 ، 192 ، 193 ، 194 ، 195 ، 196 ، 197 ، 198 ، 199.

لقد حددنا العشيريات وهي ملونة بالأخضر وعددها 9 ، كما أنها تقسم البيانات إلى 10 أقسام متساوية كل قسم يحتوي على 19 عدد، أما المئويات فقد تم تلوينها بالأحمر وعددها 99 ، كما أنها تقسم البيانات إلى 100 قسم كل قسم يحتوي على عدد واحد، وبالنسبة للقيم التي تم تلوينها باللونين فهي عشير وفي نفس الوقت مئوي كما هو الحال بالنسبة للعشير الأول الذي يوافق العدد 20 وهو في نفس الوقت المئوي العاشر.

➢ تحديد العشير الخامس والمئوي السبعون حسابيا:

تحديد العشير الخامس:

نلاحظ أن عدد بيانات السلسلة فردي (199)، وهي مرتبة تصاعديا، إذن العشير الخامس هو:

$$D_5 = X_{\frac{5(n+1)}{10}} = X_{\frac{5(199+1)}{10}} = X_{100} = 100$$

من خلال التمثيلات السابقة (القيم الملونة بالأخضر) نلاحظ أن العشير الخامس هو القيمة الملونة بالأخضر ذات الترتيب الخامس، وبالتالي هناك تطابق بين النتيجتين (الحسابية والملونة).

تحديد المئوي السبعون:

نلاحظ أن عدد بيانات السلسلة فردي (199)، وهي مرتبة تصاعديا، إذن المئوي السبعون هو:

$$P_{70} = X_{\frac{70(n+1)}{100}} = X_{\frac{70(199+1)}{100}} = X_{140} = 140$$

من خلال التمثيلات السابقة (القيم الملونة بالأحمر) نلاحظ أن المئوي السبعون هو القيمة الملونة بالأحمر ذات الترتيب السبعون، وبالتالي هناك تطابق بين النتيجتين (الحسابية والملونة).

III-3-5-2-2. حساب العشيريات والمئويات للبيانات المبوبة:

III-3-5-2-2-1. حساب العشيريات والمئويات لبيانات المتغير الكمي المنقطع (المنفصل):

إن إستخراج العشيريات والمئويات من بيانات المتغير الكمي المتقطع يتبع نفس الخطوات المتبعة في إستخراج الربيعيات، هذه الخطوات هي:

4- نحسب التكرارات التجميعية الصاعدة؛

5- نحدد رتبة العشير i وفق العلاقة التالية: $\frac{i\sum n_i}{10} = \frac{iN}{10}$، و نحدد رتبة المئوي i وفق العلاقة التالية: $\frac{i\sum n_i}{100} = \frac{iN}{100}$؛

6- نحدد في العمود الخاص بالتكرار المتجمع الصاعد من الأعلى قيمة تساوي رتبة العشير أو المئوي i أو أعلى منها مباشرة، وهذه القيمة تقابلها في المتغيرة قيمة العشير D_i أو المئوي P_i

مثال (3-28): نواصل مع معطيات المثال رقم (3-21)، التي تمثل عدد أيام التغيب لدى مجموعة من العمال.

الجدول رقم (3-30): توزيع مجموعة من عمال مؤسسة ما حسب عدد أيام التغيب.

عدد أيام الغياب X_i	0	1	2	3	4	5	$\sum$
عدد العمال n_i	10	8	4	2	3	3	30

المطلوب: أحسب العشير الثاني والمئوي الثمانون ؟

الحل:

لحساب العشيريات والمئويات نحسب التكرار التجميعي الصاعد.

الجدول رقم (3-31): استخدام جدول التوزيع التكراري في حساب العشيريات والمئويات

عدد أيام التغيب X_i	عدد العمال n_i	$n_i \uparrow$
0	10	10
1	8	18
2	4	22
3	2	24
4	3	27
5	3	30
$\sum$	30	-

- حساب العشير الثاني (D_2) :

نحسب رتبة العشير الثاني وهي : $\frac{iN}{10} = \frac{2(30)}{10} = 6$ ، ثم نبحث في قيم العمود الخاص بالتكرار المتجمع الصاعد من القيمة الأولى (10) وهي قيمة أكبر من قيمة الرتبة وهي التي تحقق الشرط (لابد أن تكون القيمة تساوي رتبة العشير أو أكبر منها مباشرة)، وبالتالي القيمة المقابلة في المتغيرة هي قيمة العشير الثاني، أي أن : $D_2 = 0$

- حساب المئوي الثمانون (P_{80}) :

نحسب رتبة المئوي الثمانون وهي : $\frac{iN}{100} = \frac{80(30)}{100} = 24$ ، ثم نبحث في قيم العمود الخاص بالتكرار المتجمع الصاعد من القيم الأولى (10 ، 18 ، 22) وهي قيم أصغر من رتبة المئوي الثمانون ، وبالتالي هي قيم مرفوضة، ، ثم ننتقل إلى القيمة الأسفل منهم (24)، وهي قيمة تساوي قيمة الرتبة وبالتالي تحقق الشرط (لابد أن تكون القيمة تساوي رتبة الربيع أو أكبر منها مباشرة)، ومنه القيمة المقابلة في المتغيرة هي قيمة المئوي الثمانون، أي أن : $P_{80} = 3$

III-3-5-2-2-2. حساب العشيريات والمئويات لبيانات المتغير الكمي المستمر (المتصل):

من أجل حساب العشيريات (D_i)، حيث $i = 1 ;2.... ;9$ أو المئويات (P_i)، حيث $i = 1 ;2.... ;99$ نتبع نفس خطوات حساب الربيعيات وهي كالتالي:

- نحسب التكرارات التجميعية الصاعدة؛
- نحدد رتبة العشير i بالعلاقة $\frac{i\sum n_i}{10} = \frac{iN}{10}$ أو رتبة المئوي i بالعلاقة $\frac{i\sum n_i}{100} = \frac{iN}{100}$؛
- نحدد في العمود الخاص بالتكرار المتجمع الصاعد من الأعلى قيمة تساوي رتبة العشير أو المئوي i أو أعلى منها مباشرة، وهذه القيمة تقابلها في المتغيرة فئة العشير D_i أو المئوي P_i .
- نحسب قيمة العشير i وفق القاعدة التالية:

$$D_i = L_1 + \frac{\frac{iN}{10} - N_0}{n_{D_i}} \times k$$

- نحسب قيمة المئوي i وفق القاعدة التالية:

$$P_i = L_1 + \frac{\frac{iN}{100} - N_0}{n_{P_i}} \times k$$

حيث : L_1 هو الحد الأدنى للفئة العشيرية أو المئوية؛

N يمثل مجموع التكرارات $N = \sum_{i=1}^{k} n_i$

N_0 هو التكرار المتجمع الصاعد للفئة قبل العشيرية أو المئوية؛

N_{Di} هو تكرار الفئة العشيرية

N_{Pi} هو تكرار الفئة المئوية

k يمثل طول الفئة العشيرية أو المئوية

مثال (3–29): نواصل مع معطيات المثال السابق رقم (3–06)

الجدول رقم (3–32): توزيع مجموعة من الأفراد حسب دخلهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	$]10-20]$	$]20-30]$	$]30-40]$	$]40-50]$	$]50-60]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: أحسب العشير الخامس والمئوي الخامس والسبعون لهذه البيانات، ثم قارن بينها وبين نتائج الربيعيات السابقة ؟

الحل: باستخدام جدول التوزيع التكراري نحسب التكرار التجميعي الصاعد.

الجدول رقم (3–33): حساب التكرار التجميعي الصاعد اعتمادا على جدول التوزيع التكراري.

فئات الدخل الشهري X_i	عدد الأفراد n_i	التكرار التجميعي الصاعد $n_i \uparrow$
$]10-20]$	5	5
$]20-30]$	8	13
$]30-40]$	12	25
$]40-50]$	25	50
$]50-60]$	10	60
المجموع	60	/

- حساب العشير الخامس (D_5):

➢ نحسب رتبة العشير الخامس وفق القاعدة : $\frac{iN}{10} = \frac{5(60)}{10} = 30$، ثم نبحث في قيم التكرار التجميعي الصاعد بداية من القيمة الأولى عن قيمة تساوي هذه الرتبة أو أكبر منها مباشرة، وبالتالي القيم الأولى (5 و 13 و 25) أقل من الرتبة ولا يمكن أن نختار أي منهم، ونلاحظ أن رتبة العشير الخامس تدخل ضمن القيمة (50) الموالية، وبالتالي الفئة المقابلة لذلك هي فئة العشير الخامس $]40-50]$.

➢ بعد تحديد فئة الربيع الثاني نحسب قيمة هذا الأخير وفق القاعدة السابقة:

$$D_5 = L_1 + \frac{\frac{iN}{10} - N_0}{n_{D_5}} \times k = 40 + \frac{\frac{5(60)}{10} - 25}{25} \times 10 = 40 + \frac{30-25}{25} \times 10 = 42$$

إذن نلاحظ أن : $D_5 = Q_2$

– حساب المئوي الخامس والسبعون (P_{75}):

➢ نحسب رتبة المئوي الخامس والسبعون وفق القاعدة : $\frac{iN}{100} = \frac{75(60)}{100} = 45$، ثم نبحث في قيم التكرار التجميعي الصاعد بداية من القيمة الأولى عن قيمة تساوي هذه الرتبة أو أكبر منها مباشرة، وبالتالي القيم الأولى (5 و 13 و 25) أقل من الرتبة ولا يمكن أن نختار أي منهم، ونلاحظ أن رتبة المئوي الخامس والسبعون تدخل ضمن القيمة (50) الموالية، وبالتالي الفئة المقابلة لذلك هي فئة المئوي الخامس والسبعون $]50-40]$ وهي نفسها فئة العشير الخامس لهذه البيانات.

➢ بعد تحديد فئة المئوي الخامس والسبعون نحسب قيمة هذا الأخير وفق القاعدة السابقة:

$$P_{75} = L_1 + \frac{\frac{iN}{100} - N_0}{n_{P_{75}}} \times k = 40 + \frac{\frac{75(60)}{100} - 25}{25} \times 10 = 40 + \frac{45-25}{25} \times 10 = 48$$

إذن نلاحظ أن : $P_{75} = D_3$

III-3-5-3. استخراج الربيعيات والعشيريات والمئويات بيانيا:

إن إستخراج الربيعيات والعشيريات والمئويات بيانيا يتم وفق طريقة استخراج الوسيط، حيث يتم تحديد رسم التكرار التجميعي الصاعد ، ثم تحدد رتبة الربيع أو العشير أو المئوي في محور التكرار التجميعي الصاعد (المحور العمودي) ويتم إسقاطها أفقيا حتى منحنى التكرار التجميعي الصاعد، وعند الالتقاء يتم إسقاط نقطة الالتقاء عموديا على محور الفئات، ونقطة السقوط على محور الفئات هي قيمة الربيع أو العشير أو المئوي.

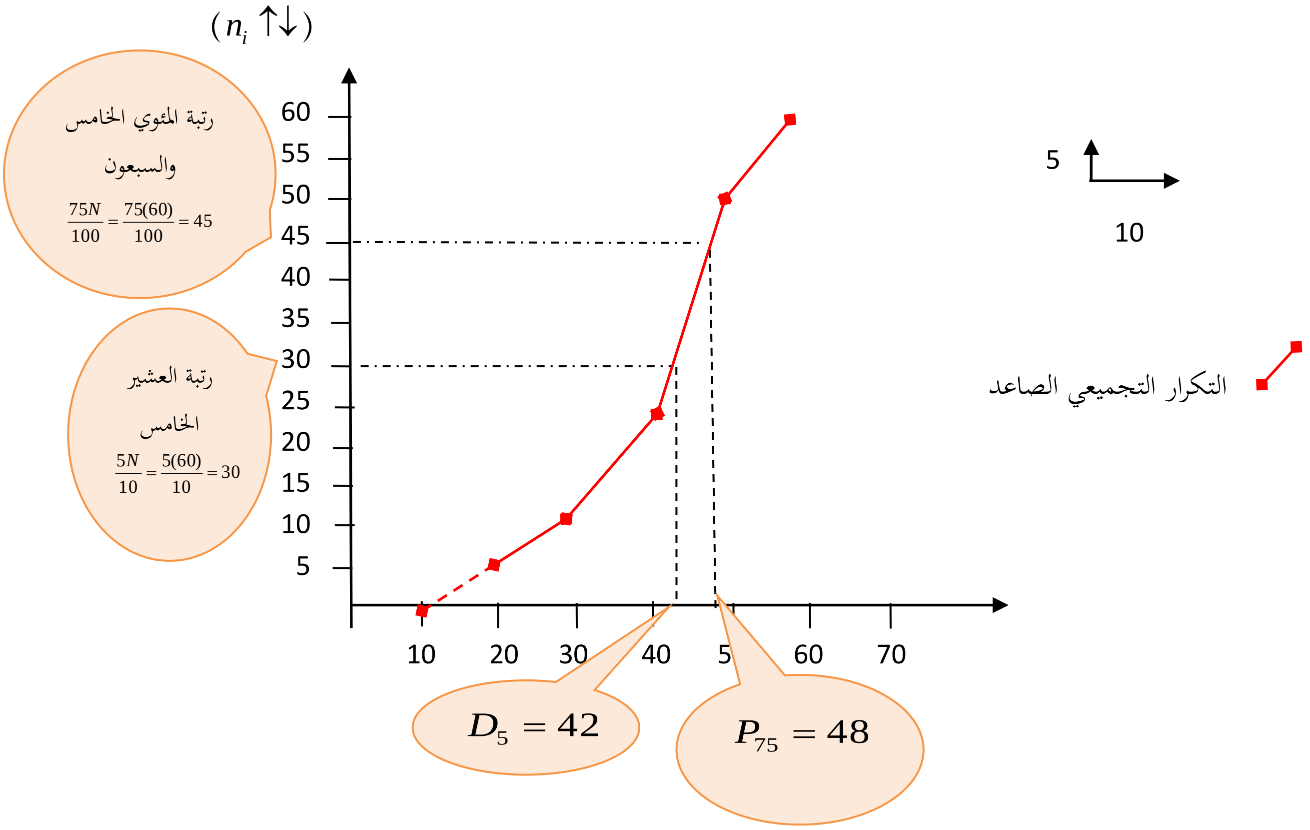
مثال (3-30): نواصل مع معطيات المثال السابق رقم (3-06)، حيث قمنا بحساب التكرار التجميعي الصاعد

الجدول رقم (3-34): حساب التكرار التجميعي الصاعد اعتمادا على جدول التوزيع التكراري.

فئات الدخل الشهري X_i	عدد الأفراد n_i	التكرار التجميعي الصاعد $n_i \uparrow$
$]20-10]$	5	5
$]30-20]$	8	13
$]40-30]$	12	25
$]50-40]$	25	50
$]60-50]$	10	60
المجموع	60	/

المطلوب: استخرج العشير الخامس والمئوي الخامس والسبعون بيانيا ، مع التأكد من ذلك بالرجوع إلى نتائج حل المثال السابق؟

الحل: نقوم برسم التكرار التجميعي الصاعد ثم نستخرج منه المطلوب



III-3-6. المنوال (The Mode):

تعتبر المنوال أحد مقاييس النزعة المركزية الهامة، لأنه المقياس الوحيد الذي يستخدم لحساب المتوسط لظاهرة ما لا يمكن قياسها بالقياس الكمي ونعني بذلك المتغيرات النوعية، فهذا المقياس يمكن إستخدامه للقيم الكمية والنوعية (الوصفية)[1] ، ويرمز له بالرمز M_0.

III-3-6-1. حساب المنوال للبيانات الأولية (غير المبوبة):

يعرف المنوال لمجموعة من البيانات بأنه القيمة أو الصفة الأكثر تكرارا أو شيوعا.[2]

مثال (3-31): تمثل المعطيات الموالية علامات 10 طلاب في مقياس الإحصاء

10 ، 12 ، 12 ، 15 ، 14 ، 19 ، 16 ، 12 ، 10 ، 13

المطلوب: حدد منوال هذه البيانات ؟

الحل: نلاحظ أن القيمة الأكثر تكرار هي العلامة **12** (تكراراتها 3)، وبالتالي : $M_O=12$

مثال (3-32): لتكن ألوان عيون مجموعة الأطفال التالية

سوداء ، سوداء ، بنية ، سوداء ، زرقاء ، زرقاء ، بنية ، سوداء ، بنية ، سوداء ، سوداء ، زرقاء ، زرقاء

المطلوب: حدد منوال هذه البيانات ؟

---

[1] - عدنان عباس حميدان وآخرون، مرجع سابق، ص 149.

[2] - نفس المرجع، ص 149.

الحل: نلاحظ أن الصفة الأكثر تكرار هي **سوداء** (تكراراتها 6)، وبالتالي : سوداء $= M_O$

III-3-6-2. حساب المنوال للبيانات المبوبة:

نظرا لوجود نوعين من المتغير الكمي ، المنقطع والمستمر فإننا سنعرض طريقة حساب المنوال حسب كل نوع.

III-3-6-2-1. حساب المنوال لبيانات المتغير الكمي المتقطع والمتغير النوعي:

إن المنوال هو الصفة المقابلة لأكبر تكرار، أو هو قيمة المتغيرة $\mathbf{x_i}$ المقابلة لأكبر تكرار في جدول التوزيع الإحصائي.

مثال (3-33): لتكن البيانات الموالية التي تمثل عدد التلاميذ في عدد من أقسام مدرسة ابتدائية

الجدول رقم (3-35): توزيع التلاميذ حسب أقسام المدرسة

القسم X_i	1	2	3	4	5
عدد التلاميذ n_i	30	25	29	32	24

المطلوب: ما هو القسم الشائع في هذه المدرسة ؟

الحل:

من خلال الجدول نلاحظ أن أكبر تكرار هو 32 وعليه فإن قيمة المنوال هي : $M_O = 4$

مثال (3-34): لتكن البيانات الموالية التي تمثل تقديرات النجاح لعدد من الطلبة.

الجدول رقم (3-36): توزيع الطلبة حسب تقديرات النجاح

تقديرات النجاح X_i	مقبول	قريب من الحسن	حسن	جيد	ممتاز
عدد الطلبة n_i	10	12	9	4	2

المطلوب: ما هو تقدير النجاح الشائع لدى الطلبة ؟

الحل:

من خلال الجدول نلاحظ أن أكبر تكرار هو 12، وعليه فإن المنوال الذي يمثل التقدير الشائع هو : قريب من الحسن$= M_O$

III-3-6-2-2. حساب المنوال لبيانات المتغير الكمي المتصل (المستمر):

يستخدم في حساب المنوال لبيانات مبوبة ذات فئات (متغير كمي متصل) عدة طرق مثل طريقة مركز الفئة وطريقة الرافعة، لكن أشهر طريقة استخداما هي طريقة الفروقات التي سنستخدمها في هذه المحاضرات، لكن قبل إستخدامها لابد من معرفة هل أن الفئات متساوية أم لا ؟.

عندما تكون أطوال الفئات متساوية نستخدم التكرارات المطلقة (العادية)، أما إذا كانت أطوال الفئات غير متساوية نقوم بتعديل التكرارات المطلقة لتصبح لدينا تكرارات معدلة تستخدم في حساب المنوال.

III-3-6-2-2-1. حساب المنوال لبيانات المتغير الكمي المتصل في حالة فئات متساوية الطول:

إذا كانت أطوال الفئات متساوية فإن نستخدم التكرارات المطلقة (الموجودة)، وأول خطوة هي تحديد الفئة المنوالية، وهي الفئة المقابلة لأكبر تكرار n_i والتي ينتمي إليها المنوال، ثم يتم حساب المنوال باستخدام العلاقة التالية:

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times k$$

حيث :

L_1 : الحد الأدنى للفئة المنوالية

Δ_1 : الفرق بين تكرار الفئة المنوالية و تكرار الفئة السابقة لها

Δ_2 : الفرق بين تكرار الفئة المنوالية و تكرار الفئة اللاحقة لها

k : طول الفئة المنوالية

مثال (3–35): نعود إلى معطيات المثال السابق رقم (3–06) والتي تمثل توزيع مجموعة من الأفراد حسب دخلهم الشهري.

الجدول رقم (3–37): توزيع مجموعة من الأفراد حسب دخلهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: أحسب المنوال ؟

الحل: نلاحظ أن أطوال الفئات متساوية وبالتالي نستخدم التكرارات المطلقة.

الجدول رقم (3–38): استخدام جدول التوزيع التكراري في حساب المنوال

فئات الدخل الشهري X_i	عدد الأفراد n_i
$]20-10]$	5
$]30-20]$	8
$]40-30]$	12
$]50-40]$	25
$]60-50]$	10
المجموع	60

- نقوم أولا بتحديد الفئة المنوالية، وهي الفئة المقابلة لأكبر تكرار(25) إذن الفئة المنوالية هي $]50-40]$.
- نحسب المنوال باستخدام العلاقة التالية:

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times k = 40 + \frac{(25-12)}{(25-12)+(25-10)} \times 10 = 40 + \frac{13}{13+15} \times 10 = 44{,}64 \times 10^3 DA$$

III–3–6–2–2–1. حساب المنوال لبيانات المتغير الكمي المتصل في حالة فئات غير متساوية الطول:

من أجل حساب المنوال للبيانات الخاصة بمتغير كمي مستمر في حالة فئات غير متساوية الطول نقوم أولا بتعديل التكرارات وفق قاعدة $n_i^* = \frac{n_i}{L_i} \times L^*$ ، مع L^* هي طول الفئة المختار ، ومن الأفضل أن يكون أصغر طول للفئات، ثم يتم تحديد الفئة المنوالية،

وهي الفئة المقابلة لأكبر تكرار معدّل n_i^*، ويتم حساب المنوال باستخدام العلاقة السابقة، لكن باستخدام التكرار المعدّل بدل التكرار المطلق :

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times k$$

حيث :

L_1 : الحد الأدنى للفئة المنوالية

Δ_1 : الفرق بين التكرار المعدل للفئة المنوالية و التكرار المعدل للفئة السابقة لها

Δ_2 : الفرق بين التكرار المعدل للفئة المنوالية و التكرار المعدل للفئة اللاحقة لها

k : طول الفئة المنوالية

مثال (3–36): تمثل البيانات الموالية الاستهلاك الأسري لعدد من العائلات خلال شهر ما.

الجدول رقم (3–39): توزيع مجموعة من العائلات حسب إستهلاكهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	$]20-10]$	$]26-20]$	$]40-26]$	$]50-40]$	$]70-50]$	المجموع
عدد الأفراد n_i	2	10	14	6	8	40

المطلوب: أحسب المنوال ؟

الحل: نلاحظ أن أطوال الفئات غير متساوية وبالتالي نقوم بحساب التكرارات المعدلة من أجل إستخدامها في حساب المنوال.

الجدول رقم (3–40): استخدام جدول التوزيع التكراري في تعديل التكرارات

فئات الدخل الشهري X_i	عدد الأفراد n_i	طول الفئة L_i	التكرار المعدل n_i^*
$]20-10]$	2	10	$n_1^* = \frac{2}{10} \times 6 = 1{,}2$
$]26-20]$	10	6	$n_2^* = \frac{10}{6} \times 6 = 10$
$]40-26]$	14	14	$n_3^* = \frac{14}{14} \times 6 = 6$
$]50-40]$	6	10	$n_4^* = \frac{6}{10} \times 6 = 3{,}6$
$]70-50]$	8	20	$n_5^* = \frac{8}{20} \times 6 = 2{,}4$
المجموع	40	/	/

- نقوم بتحديد الفئة المنوالية، وهي الفئة المقابلة لأكبر تكرار معدل (10) إذن الفئة المنوالية هي $]26-20]$.
- نحسب المنوال باستخدام العلاقة السابقة:

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times k = 20 + \frac{(10-1{,}2)}{(10-1{,}2)+(10-6)} \times 6 = 20 + \frac{8{,}8}{8{,}8+4} \times 6 = 24{,}125 \times 10^3 \, DA$$

III-3-6-3. استخراج المنوال بيانيا:

إن المنوال كما هو الحال بالنسبة للوسيط يمكن تحديده بيانيا، ويكون ذلك من خلال المدرج التكراري، حيث يتم استخراجه من الفئة المنوالية (الفئة الأطول)، وهذا بإيصال نقطة التقاء نهاية الفئة قبل المنوالية مع بداية المستطيل للفئة المنوالية بنهاية الفئة المنوالية من الأعلى (القمة) وهي قطعة المستقيم الأولى، أما قطعة المستقيم الثانية فتحدد ببداية الفئة المنوالية من الأعلى بنقطة التقاء نهاية الفئة المنوالية مع بداية المستطيل للفئة اللاحقة لها من الأعلى، وعند نقطة تقاطع القطعتين المستقيمتين المرسومتين نسقط عمود على المحور الأفقي، فيمسه عند نقطة تعبر عن قيمة المنوال، ولتوضيح كيفية تحديد المنوال بيانيا نستعين بمعطيات المثال السابق رقم (3-06).

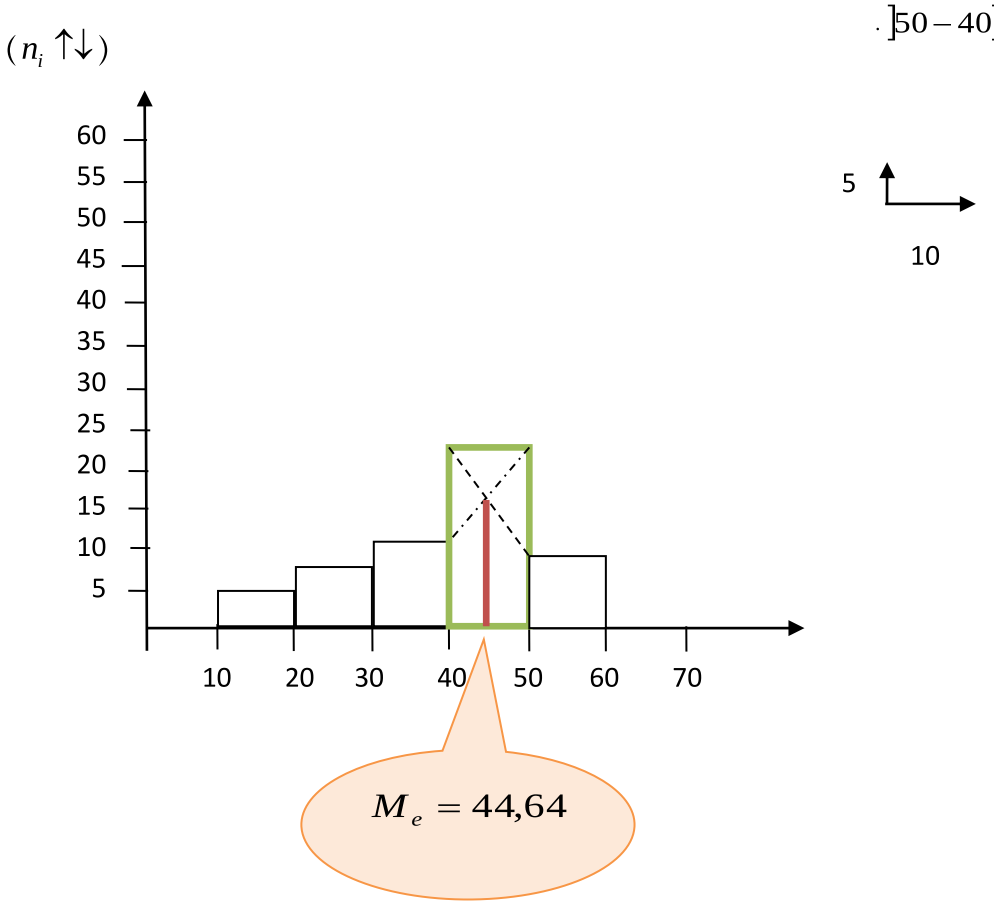
مثال (3-37): نعود إلى معطيات المثال السابق رقم (3-06) التي تمثل توزيع مجموعة من الأفراد حسب دخلهم الشهري.

الجدول رقم (3-41): توزيع مجموعة من الأفراد حسب دخلهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: إستخرج المنوال بيانيا ؟

الحل: نلاحظ أن أطوال الفئات متساوية وبالتالي نستخدم التكرارات المطلقة في رسم المدرج التكراري، كما أن الفئة المنوالية هي الفئة المقابلة لأكبر تكرار $]50-40]$.



III-3-6-4. خواص المنوال :

يتميز المنوال بمجموعة خصائص قد نعتبرها محاسن، ومن أبرزها: [1]

- عدم تأثره بالقيم المتطرفة (الشاذة)؛
- يمثل غالبية المشاهدات مع أن حسابه لا يحتاج لجميع قيم التوزيع؛
- إمكانية حسابه في حالة الجداول التكرارية ذات الفئات المفتوحة؛
- له فائدة كبيرة للبيانات الوصفية إذا أردنا الحصول على مقياس للنزعة المركزية لهذه البيانات؛
- يمكن تحديده بيانيا.

III-3-7. العلاقة التقريبية بين المتوسط الحسابي والوسيط والمنوال:

توجد علاقة رياضية بين أهم مقاييس النزعة المركزية والمتمثلة في المتوسط الحسابي، الوسيط والمنوال تطبق خاصة في البيانات المبوبة لمتغير كمي مستمر (فئات)، والتي تحتوي على منوال واحد، هذه العلاقة الرياضية تتحدد وفق ما يلي: [2]

إذا كان لمجموعة البيانات منوال واحد فإن المتوسط والوسيط والمنوال تربطهم إحدى العلاقات التالية:

- إذا كان التوزيع التكراري متماثلا فإن قيم كل من المتوسط الحسابي والوسيط والمنوال متساوية، أي $\overline{X} = M_e = M_0$ ؛
- إذا كان التوزيع التكراري قريبا من التماثل كانت المقاييس الثلاثة متقاربة، وكلما إبتعد التوزيع عن التماثل كلما تباعدت المقاييس عن بعضها البعض، وقد وجد في حالة التوزيعات التكرارية القريبة من التماثل علاقة تقريبية تكون صحيحة هي $\overline{X} - M_0 = 3(\overline{X} - M_e)$ ، ومن خلال هذه الطريقة يمكن إيجاد المتوسط الحسابي من الجداول التكرارية المفتوحة بعد معرفة الوسيط والمنوال.

مثال (3-38): نعود إلى معطيات المثال السابق رقم (3-06) التي تمثل توزيع مجموعة من الأفراد حسب دخلهم الشهري.

الجدول رقم (3-42): توزيع مجموعة من الأفراد حسب دخلهم الشهري. (الوحدة: 10^3 دج)

فئات الدخل الشهري x_i	أقل من 20	$[20-30[$	$[30-40[$	$[40-50[$	$[50-60[$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: لنفرض أن الفئة الأولى مفتوحة كما هو موضح في الجدول، وبناءا على قيم الوسيط والمنوال التي تم حسابها سابقا قم بحساب المتوسط الحسابي، ثم قارنه مع قيمته التي تم حسابها سابقا ؟

الحل: لنستخدم العلاقة التالية في حساب المتوسط الحسابي من البيانات ذات الفئة المفتوحة.

---

[1] - عدنان عباس حميدان وآخرون، مرجع سابق، ص 156

[2] - خالد أحمد فرحان المشهداني ورائد عبد الخالق عبد الله العبيدي ، مبادئ الإحصاء – متضمن التحليل الاحصائي SPSS ، دار الأيام ، عمان ، الأردن، ، 2013، ص ص : 78 ، 79.

$$\overline{X} - M_0 = 3(\overline{X} - M_e)$$
$$\overline{X} = 3\overline{X} - 3M_e + M_0$$
$$2\overline{X} = 3M_e - M_0$$
$$\overline{X} = \frac{3M_e - M_0}{2}$$

قمنا بحساب الوسيط والمنوال لنفس البيانات ووجدنا أن : $M_e = 42$ و $M_0 = 44{,}64$ ، وبالرغم من أن القيمتين غير متساويتين ، أي عدم تماثل التوزيع التكراري إلا أننا نفترضه قريب من التماثل ونستخدم العلاقة السابقة في حساب المتوسط الحسابي كالتالي:

$$\overline{X} = \frac{3M_e - M_0}{2} = \frac{3(42) - 44{,}64}{2} = 40{,}68$$

لقد وجدنا سابقا أن المتوسط الحسابي لنفس البيانات هو $\overline{X} = 39{,}5$ ، وهي قيمة قريبة من المحسوبة بواسطة العلاقة التي تربط بين المقاييس الثلاثة، وكلما كان التوزيع أكثر تماثلا كانت القيم أقرب إلى التساوي.

المحور الرابع:

مقاييس التشتت

IV-1. تمهيد:

إن مقاييس النزعة المركزية تعطينا مؤشرات إحصائية ذات دلالة وصفية لتوضيح الشكل العام للتوزيع البياني دون الإشارة إلى ماهية التوزيع وعما يجري داخل هذا التوزيع من تناثر أو تباعد أو اقتراب مفردات التوزيع عن بعضها البعض، فقد يتساوى المتوسط الحسابي لمجموعتين من البيانات، وتعطينا نظرة أولى أن هذه البيانات متشابهة، لكن عند فحصها نجد هناك تباين بين هذه البيانات، فقد تكون الأولى أكثر تقارب بالمقارنة مع الثانية، وبالتالي يمكن القول أن مقاييس النزعة المركزية لا تعطي صورة شاملة عن تموقع البيانات فيما بينها، ولمعرفة أشمل وأوسع عن هذه البيانات يتطلب استخدام مقاييس أخرى تكمل ما توصلت إليه مقاييس النزعة المركزية، من بين هذه المقاييس مقاييس التشتت، التي لها مجموعة من الافادات هي: [1]

- تزودنا بمعلومات حول تبعثر (تشتت)، أو تجمع البيانات داخل التوزيع، وحول المتوسط الحسابي لهذا التوزيع؛
- تقيم تلك المقاييس مدى فعالية مقاييس النزعة المركزية، فمثلا كلما تجمعت القيم حول متوسطها الحسابي كان هذا الأخير يمثل بشكل جيد هذه القيم والعكس صحيح، بمعنى أنه كلما كان مقياس التشتت صغيرا كان هذا مؤشرا على أن مقياس النزعة المركزية يمثل بياناته أصدق تمثيل؛
- لمقاييس التشتت دور هام من الناحية التحليلية فيما يخص الاستدلال الاحصائي، فالعينات تتم دراستها للاستدلال على المجتمعات التي سحبت منها، ففي كل الأحوال يستخدم المتوسط الحسابي للعينة للاستدلال على الوسط الحسابي للمجتمع الاحصائي الذي سحبت منه العينة، لكن في غالب الأحيان لا ينطبق الوسط الحسابي للعينة مع المتوسط الحسابي للمجتمع، فهناك خطأ متوقع نتيجة لاستخدام أسلوب العينات، وهنا يأتي دور مقاييس التشتت في المساهمة بمعلومات توضح حجم هذا الخطأ؛
- إن مقاييس التشتت لها الدور التكميلي لمقاييس النزعة المركزية في وصف التوزيع البياني، ولها دور هام أيضا في المساعدة على مقارنة توزيع بياني مع آخر.

إذن يمكن تعريف التشتت على أنه " مدى تباعد مجموعة القيم عن بعضها البعض أو عن القيمة التي تمثل مركز تلك المجموعة " [2]، ويقاس التشتت بمجموعة من المقاييس يمكن تقسيمها إلى قسمين مقاييس التشتت المطلقة والأخرى نسبية.

IV-2. قياس التشتت أو الانتشار Variability or dispersion :

هناك قسمين من مقاييس التشتت، مقاييس التشتت المطلقة والتي تكون وحدة قياسها هي وحدة قياس القيم الأصلية، حيث أن البعض منها يقيس تقارب أو تباعد القيم عن بعضها البعض وهي المدى والانحراف الربيعي، والأخرى تقيس قرب أو بعد القيم من قيمة معينة كالمتوسط الحسابي وهي الانحراف المتوسط والتباين، أما القسم الثاني فهي مقاييس التشتت النسبية والتي تكون خالية من وحدات القياس بل تكون كنسبة مئوية، وهي معامل الاختلاف النسبي ومعامل الاختلاف الربيعي.

[1] - عدنان عباس حميدان وآخرون، مرجع سابق، ص ص : 182 ، 183.

[2] - محمد راتول، مرجع سابق، ص 138.

IV-2-1. مقاييس التشتت المطلقة:

IV-2-1-1. المدى المطلق (العام) The Range:

هو أبسط مقياس من مقاييس التشتت، يستخدم لما يكون الهدف هو الحصول على قياس تشتت سريع وبسيط، ويرمز له بالرمز R ، أما طريقة حسابه فتختلف حسب نوع البيانات.

IV-2-1-1-1. المدى المطلق في حالة بيانات أولية (غير مبوبة):

إن المدى المطلق أو العام في حالة بيانات غير مبوبة هو عبارة عن الفرق بين أكبر قيمة في البيانات وأصغر قيمة لها، أي:

المدى المطلق (العام) = أكبر قيمة – أصغر قيمة $$R = X_{max} - X_{min}$$

مثال (4-01): لتكن البيانات الموالية تمثل علامات بعض طلبة فوجين من قسم السنة الأولى في مقياس الإحصاء

الفوج الأول: 9 ، 8 ، 12 ، 17 ، 12 ، 19 ، 11 ، 5 ، 18 ، 14 ، 12 ، 11 ، 10 ، 9

الفوج الثاني: 8 ، 9 ، 9 ، 9 ، 7 ، 2 ، 10 ، 8 ، 7 ، 7 ، 9 ، 16 ، 8 ، 9 ، 10

المطلوب: أحسب المدى المطلق لعلامات الطلبة في كلا الفوجين، ثم قارن بينهما؟

الحل: لدينا

$$R_1 = X_{max} - X_{min} = 19 - 5 = 14$$

$$R_2 = X_{max} - X_{min} = 16 - 2 = 14$$

نلاحظ أن المديين المطلقين متساويين، لكن بعد فحص البيانات للفوجين وجدنا أن علامات طلبة الفوج الأول أكثر إنتشار بالمقارنة مع علامات طلبة الفوج الثاني، التي بالرغم من أنها أكثر تجانس إلا أنها تحتوي على علامتين شاذتين أدتا إلى زيادة قيمة التشتت.

IV-2-1-1-2. المدى المطلق (العام) في حالة بيانات مبوبة:

إن حساب المدى المطلق من البيانات المبوبة يختلف حسب نوع البيانات، حيث :

- إذا كانت البيانات المبوبة خاصة بمتغير كمي منفصل (متقطع) فإن المدى المطلق يحسب وفق العلاقة :

المدى المطلق = القيمة النقطية الكبرى للمتغيرة – القيمة النقطية الصغرى للمتغيرة

$$R = X_{max} - X_{min}$$

- إذا كانت البيانات المبوبة خاصة بمتغير كمي متصل (مستمر) فإن المدى المطلق يحسب وفق العلاقة :

المدى المطلق = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

$$R = L_{max\ k} - L_{min\ 1}$$

مثال (4-02): لتكن البيانات المبوبة التالية الخاصة بالجدول التكراري المزدوج رقم (2-10) السابق

الجدول رقم (4–01): جدول مزدوج لبيانات عدد من الأسر حول عدد أطفالها وحجم نفقاتها.

النفقات / عدد الأطفال	]25 20]	]30 25]	]35 30]	]40 35]	]45 40]	]50 45]	المجموع
1	4	1	0	0	0	0	5
2	0	1	3	0	0	0	4
3	0	0	2	3	0	0	5
4	0	0	0	1	5	0	6
5	0	0	0	0	0	4	4
المجموع	4	2	5	4	5	4	24

المطلوب: أحسب المدى المطلق للمتغيرتين؟

الحل: لدينا متغيرة كمية منفصلة وأخرى متصلة

– المدى المطلق (العام) لمتغيرة عدد الأطفال (كمية منفصلة):

$$R = X_{\max} - X_{\min} = 5 - 1 = 4$$

– المدى المطلق (العام) لمتغيرة النفقات (كمية متصلة):

$$R = L_{\max\ k} - L_{\min\ 1} = 50 - 20 = 30$$

IV–2–1–1–3. خواص المدى المطلق (العام):

يتميز المدى المطلق (العام) بعدد من الخصائص أبرزها:

- من أبسط مقاييس التشتت حسابيا لأنه يعتمد في ذلك على قيمتين فقط؛
- من عيوبه أنه شديد التأثر بالقيم الشاذة (المتطرفة)؛
- بما أنه يعتمد على حدود الفئات فلا يمكن حسابه من جداول التوزيع التكراري المفتوحة؛
- لا يصلح إلا للمقارنة بين توزيعين لمتغير واحد ولهما نفس وحدة القياس.

IV–2–1–2. المدى الربيعي (Inter-quartile range):

نظرا لعدم القدرة على حساب المدى المطلق (العام) في حالة البيانات ذات الفئات المفتوحة، وكذلك تأثره بالقيم المتطرفة، والتي من خلالها يصبح المدى المطلق غير مجد، وجد مقياس تشتت آخر لمعالجة هذه العيوب يسمى بالمدى الربيعي، هذا الأخير يهمل الحدود العليا (القيم الشاذة والفئات المفتوحة) الموجودة في الربعين الأول والأخير، ويعتمد على القيم الموجودة بين الربيعيين الأول والثالث، ويرمز لهذا المدى الربيعي بـ I_Q.

IV–2–1–2–1. قياس المدى الربيعي:

إن قياس المدى الربيعي يعتمد على الربيعيين الأول والثالث، اللذين عرضنا سابقا طرق حسابهما سواءا كانت البيانات أولية (غير مبوبة) أو مبوبة، وبالتالي نكتفي بإعطاء صياغته الرياضية، حيث :

المدى الربيعي = الربيع الثالث – الربيع الأول $$I_Q = Q_3 - Q_1$$

مثال (4–03): نظرا لإعتماد المدى الربيعي على الربيعيين الأول والثالث، وقد تطرقنا سابقا (المحور الثالث) لطرق حسابهما من البيانات الأولية والمبوبة، فإننا سنكتفي بالاعتماد فقط على أحد حلول الأمثلة السابقة وليكن المثال رقم (3-26)، حي توصلنا إلى أن $Q_1 = 31{,}66$ و $Q_3 = 48$.

المطلوب: أحسب المدى الربيعي، ثم قارنه مع $I_{Q_A} = 10$ ، وأيهما أفضل؟

الحل:

$$I_Q = Q_3 - Q_1 = 48 - 31{,}66 = 16{,}34$$

نلاحظ أن : $I_{Q_A} < I_Q$ ، أي التشتت المعطى أقل من التشتت المحسوب، وبالتالي تشتت البيانات الأفضل هو التشتت الأقل قيمة.

IV–2–1–2–2. تحديد المدى الربيعي بيانيا:

لقد تطرقنا سابقا إلى طريقة استخراج الربيعيات بيانيا، وبالتالي انطلاقا من هذا التمثيل يمكن استخراج المدى الربيعي.

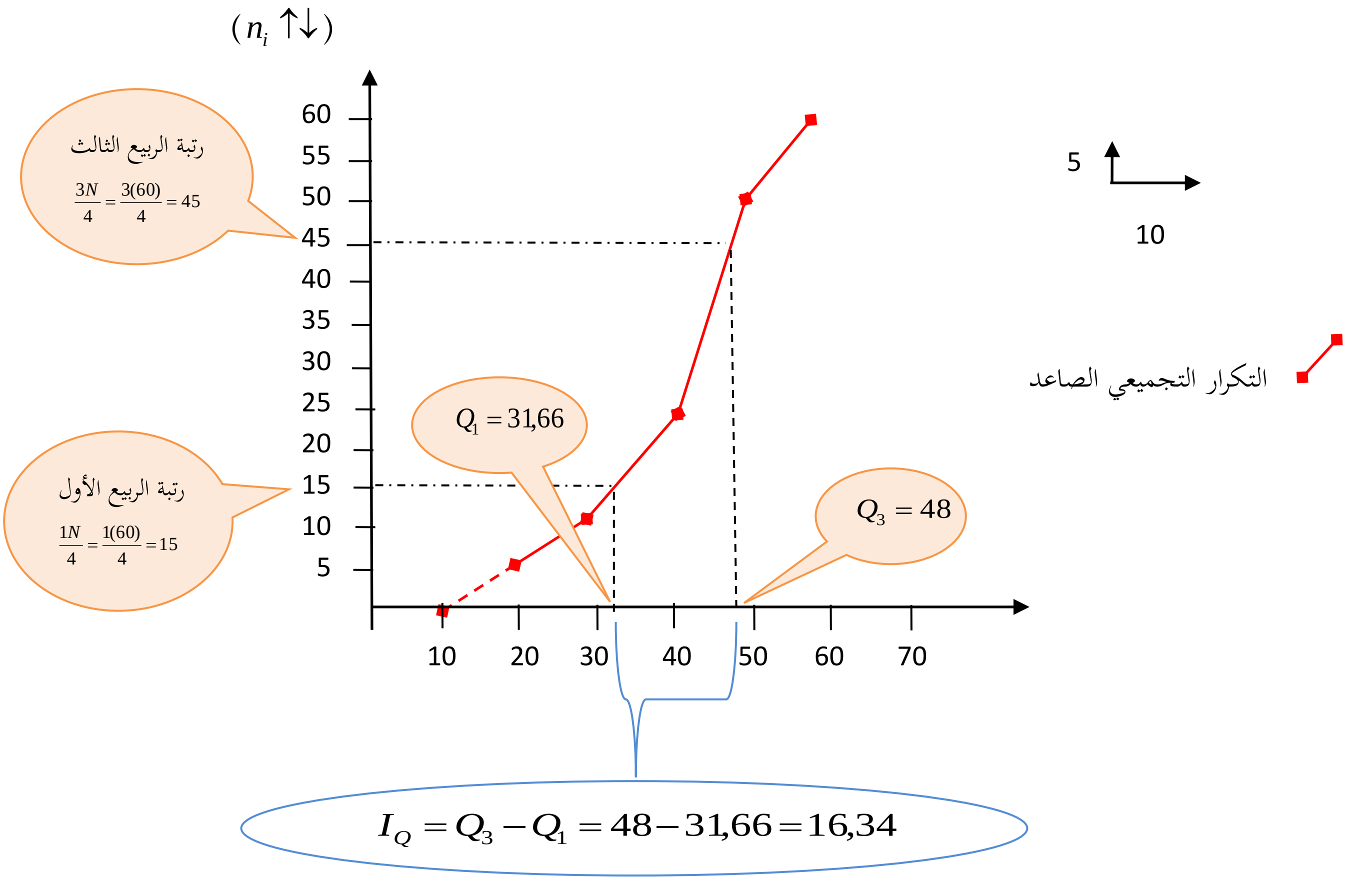
مثال (4–04): من خلال معطيات المثال (3-06) المبينة في الجدول الموالي والحل النموذجي للمثال رقم (3-26).

الجدول رقم (4–02): الجدول التكراري الذي يمثل توزيع مجموعة من الأفراد حسب دخلهم الشهري (الوحدة: 10^3 دج)

فئات الدخل الشهري X_i	عدد الأفراد n_i	التكرار التجميعي الصاعد $n_i \uparrow$
$]20-10]$	5	5
$]30-20]$	8	13
$]40-30]$	12	25
$]50-40]$	25	50
$]60-50]$	10	60
المجموع	60	/

المطلوب: حدد الربيعيين الأول والثالث بيانيا، ثم حدد المدى الربيعي؟

الحل : لقد توصلنا من خلال حل معطيات هذا المثال أن رتبة الربيع الأول هي : $\frac{iN}{4} = \frac{1(60)}{4} = 15$ ، وأن رتبة الربيع الثالث هي $\frac{iN}{4} = \frac{3(60)}{4} = 45$ ، إذن تحديد الربيعيين الأول والثالث بيانيا يكون كما يلي:



IV-2-1-2-3. خصائص المدى الربيعي:

يتميز المدى الربيعي بالخصائص التالية:[1]

- يضم 50% من المجتمع مهما كان التوزيع الاحصائي؛
- يتغير طوله مقارنة بالمدى العام حسب طبيعة المجتمع؛
- إستعماله محدود نظرا لبساطته، غير أنه أحسن من المدى العام، لأنه يمكن حسابه من جداول التوزيعات التكرارية المفتوحة؛
- يستعمل في المقارنة بين توزيعين إحصائيين أو أكثر، ويمكن إيجاده بيانيا.

IV-2-1-3. الانحراف المتوسط:

هو مقياس من مقاييس التشتت التي تقيس مقدار الاختلاف (الانحراف) بين القيم ومتوسطها الحسابي، فإذا كان هذا المقدار كبيرا دل ذلك على تشتت البيانات (عدم تجانسها)، والعكس صحيح، ويرمز له بالرمز $E_{\overline{X}}$.

[1] - جيلالي جلاطو، مرجع سابق، ص 71.

IV-2-1-3-1. حساب الانحراف المتوسط من البيانات الأولية (غير المبوبة):

إذا كانت قيم الظاهرة المدروسة $x_1 ; x_2 ; .... ; x_n$ والتي عددها **n** فإن الانحراف المتوسط يعطى بالصيغة التالية:

$$E_{\overline{X}} = \frac{\left|x_1 - \overline{X}\right| + \left|x_2 - \overline{X}\right| + ... + \left|x_n - \overline{X}\right|}{n}$$

$$E_{\overline{X}} = \frac{\sum_{i=1}^{n}\left|x_i - \overline{X}\right|}{n}$$

مثال (4-05): لتكن بيانات المثال السابق رقم (3-01) ، والتي وجدنا أن متوسطها الحسابي $\overline{X} = 7{,}5$

2 ، 4 ، 8 ، 9 ، 10 ، 8 ، 10 ، 9 ، 10 ، 7 ، 5 ، 8

المطلوب: أحسب الانحراف المتوسط للبيانات ثم قارنه مع إنحراف متوسط لبيانات أخرى قدر بـ $E_{\overline{Y}} = 4$

الحل: لدينا

$$E_{\overline{X}} = \frac{\left|x_1 - \overline{X}\right| + \left|x_2 - \overline{X}\right| + ... + \left|x_{12} - \overline{X}\right|}{n}$$

$$E_{\overline{X}} = \frac{|2-7{,}5|+|4-7{,}5|+|8-7{,}5|+|9-7{,}5|+|10-7{,}5|+|8-7{,}5|+|10-7{,}5|+|9-7{,}5|+|10-7{,}5|+|7-7{,}5|+|5-7{,}5|+|8-7{,}5|}{12}$$

$$E_{\overline{X}} = \frac{5{,}5+3{,}5+0{,}5+1{,}5+2{,}5+0{,}5+2{,}5+1{,}5+2{,}5+0{,}5+2{,}5+0{,}5}{12}$$

$$E_{\overline{X}} = \frac{5{,}5+3{,}5+0{,}5+1{,}5+2{,}5+0{,}5+2{,}5+1{,}5+2{,}5+0{,}5+2{,}5+0{,}5}{12} = \frac{24}{12} = 2$$

$$E_{\overline{X}} = 2$$

نلاحظ أن : $E_{\overline{Y}} > E_{\overline{X}}$ أي تشتت البيانات المعطاة أقل من تشتت البيانات الأخرى، وبالتالي هي أكثر تجانس .

IV-2-1-3-2. حساب الانحراف المتوسط من البيانات المبوبة:

إذا كانت القيم النقطية للمتغير الكمي المنفصل هي : $x_1 ; x_2 ; .... ; x_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، أو إذا كانت مراكز الفئات للمتغير الكمي المستمر هي : $C_1 ; C_2 ; .... ; C_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها فإن الانحراف المتوسط لهذه البيانات يحسب وفق القاعدة التالية:

➢ **المتغير الكمي المنفصل (المتقطع):**

$$E_{\overline{X}} = \frac{\left|x_1 - \overline{X}\right|n_1 + \left|x_2 - \overline{X}\right|n_2 + ... + \left|x_k - \overline{X}\right|n_k}{\sum_{i=1}^{k} n_i}$$

$$E_{\overline{X}} = \frac{\sum_{i=1}^{k} |x_i - \overline{X}| n_i}{\sum_{i=1}^{k} n_i}$$

➢ **المتغير الكمي المتصل (المستمر):**

$$E_{\overline{X}} = \frac{|C_1 - \overline{X}| n_1 + |C_2 - \overline{X}| n_2 + ... + |C_k - \overline{X}| n_k}{\sum_{i=1}^{k} n_i}$$

$$E_{\overline{X}} = \frac{\sum_{i=1}^{k} |C_i - \overline{X}| n_i}{\sum_{i=1}^{k} n_i}$$

مثال (4-06): لتكن بيانات المثال السابق رقم (3-04) ، والتي وجدنا أن متوسطها الحسابي $\overline{X} = 2$

الجدول رقم (4-03): توزيع مجموعة عمال مؤسسة حسب عدد أيام تغيبها .

عدد أيام الغياب X_i	0	1	2	3	4	5	$\sum$
عدد العمال n_i	10	8	4	2	3	3	30

المطلوب: إيجاد الانحراف المتوسط للبيانات ؟

الحل: نستعين بجدول التوزيع التكراري

الجدول رقم (4-04): طريقة حساب الانحراف المتوسط باستخدام جدول التوزيع التكراري، حيث $\overline{X} = 2$

عدد أيام التغيب X_i	عدد العمال n_i	$\lvert x_i - \overline{X} \rvert$	$\lvert x_i - \overline{X} \rvert \ n_i$
0	10	2	20
1	8	1	8
2	4	0	0
3	2	1	2
4	3	2	6
5	3	3	9
المجموع	30	/	45

إذن انحراف المتوسط هو:

$$E_{\overline{X}} = \frac{|x_1 - \overline{X}| n_1 + |x_2 - \overline{X}| n_2 + ... + |x_k - \overline{X}| n_k}{\sum_{i=1}^{k} n_i} = \frac{45}{30} = 1{,}5$$

مثال (4–07): لتكن بيانات المثال السابق رقم (3–06) ، والتي وجدنا أن متوسطها الحسابي $\overline{X}=39{,}5$

الجدول رقم (4–05): توزيع مجموعة من الأفراد حسب دخلهم الشهري (الوحدة : 10^3دج)

فئات الدخل الشهري x_i	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: إيجاد الانحراف المتوسط لدخل الأفراد ؟

الحل: إعتمادا على جدول التوزيع التكراري يتم حساب الانحراف المتوسط وهذا بعد استخراج مراكز الفئات.

الجدول رقم (4–06): حساب الانحراف المتوسط اعتمادا على جدول التوزيع التكراري، حيث $\overline{X}=39{,}5$

فئات الدخل الشهري X_i	عدد الأفراد n_i	مراكز الفئات c_i	$\lvert x_i-\overline{X}\rvert$	$\lvert x_i-\overline{X}\rvert\ n_i$
$]20-10]$	5	15	24,5	122,5
$]30-20]$	8	25	14,5	116
$]40-30]$	12	35	4,5	54
$]50-40]$	25	45	5,5	137,5
$]60-50]$	10	55	15,5	155
المجموع	60	/	/	585

إذن الانحراف المتوسط للدخل هو:

$$E_{\overline{X}}=\frac{\sum_{i=1}^{k}\lvert C_i-\overline{X}\rvert n_i}{\sum_{i=1}^{k}n_i}=\frac{585}{60}=9{,}75$$

IV–2–1–3–3. خواص الانحراف المتوسط:

يتميز الانحراف المتوسط بالخصائص التالية: [1]

- يأخذ بعين الاعتبار جميع القيم (إنحرافات كل قيمة عن المتوسط الحسابي)، فهو يعتبر مقياسا جيدا للتشتت بالمقارنة مع مقياسي المدى والانحراف الربيعي؛
- يمكن حسابه من كل من المتوسط الحسابي والوسيط، إلا أن قيمته عند حسابه من الوسيط تكون دائما أقل من قيمته عند حسابه من المتوسط، ففي التوزيعات المتماثلة والقريبة من التماثل يفضل حسابه من المتوسط الحسابي، أما في التوزيعات غير المتماثلة فيفضل حسابه من الوسيط؛
- سهل الحساب؛

[1] - عدنان عباس حميدان، مرجع سابق، ص ص : 199 ، 200

- ومن مساوئه أنه نادر الاستعمال، ولا يمكن حسابه من الجداول التكرارية ذات الفئات المفتوحة.
- يتأثر بالقيم المتطرفة ويعتمد في حسابه على جميع القيم؛

IV-2-1-4. التباين (variance) والانحراف المعياري (Standard Deviation):

يعتبر التباين من أهم مقاييس التشتت، فهو عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي، ويرمز له بإحدى الرمزين $V(x)$ أو σ^2 ، أما الانحراف المعياري فهو جذر التباين ويرمز له بالرمز σ

IV-2-1-4-1. حساب التباين والانحراف المعياري من البيانات الأولية (غير المبوبة):

إذا كانت قيم الظاهرة المدروسة x_1 ; x_2 ; .... ; x_n والتي عددها **n** فإن تباين هذه القيم يتم استخراجه وفق الصياغة الرياضية التالية :

$$V(x) = \sigma^2 = \frac{(x_1 - \overline{X})^2 + (x_2 - \overline{X})^2 + ... + (x_n - \overline{X})^2}{n}$$

$$V(x) = \sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{X})^2}{n}$$

كما يمكن كتابة التباين بالقاعدة التالية :

$$V(x) = \sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}$$

حيث : $\overline{X} = \mu$

أما الانحراف المعياري فهو : $\sigma = \sqrt{V(x)}$

مثال (4-08): لتكن بيانات المثال السابق رقم (3-01) ، والتي وجدنا أن متوسطها الحسابي $\overline{X} = 7{,}5$

2 ، 4 ، 8 ، 9 ، 10 ، 8 ، 10 ، 9 ، 10 ، 7 ، 5 ، 8

المطلوب: أحسب الانحراف المعياري للبيانات ؟

الحل: لدينا

$$V(x) = \sigma^2 = \frac{(x_1 - \overline{X})^2 + (x_2 - \overline{X})^2 + ... + (x_n - \overline{X})^2}{n}$$

$$V(x) = \frac{(2-7{,}5)^2 + (4-7{,}5)^2 + (8-7{,}5)^2 + (9-7{,}5)^2 + (10-7{,}5)^2 + (8-7{,}5)^2 + (10-7{,}5)^2 + (9-7{,}5)^2 + (10-7{,}5)^2 + (7-7{,}5)^2 + (5-7{,}5)^2 + (8-7{,}5)^2}{12}$$

$$V(x) = \frac{30{,}25 + 12{,}25 + 0{,}25 + 2.25 + 6{,}25 + 0{,}25 + 6{,}25 + 2{,}25 + 6{,}25 + 0{,}25 + 6{,}25 + 0{,}25}{12}$$

$$V(x)=\frac{73}{12}=6{,}08$$

إذن الانحراف المعياري هو: $\sigma=\sqrt{V(x)}=\sqrt{6{,}08}=2{,}46$

IV-2-1-4-2. حساب التباين والانحراف المعياري من البيانات المبوبة:

إذا كانت القيم النقطية للمتغير الكمي المنفصل هي : $x_1 ; x_2 ; .... ; x_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، أو إذا كانت مراكز الفئات للمتغير الكمي المستمر هي : $C_1 ; C_2 ; .... ; C_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها فإن التباين يحسب وفق القاعدة التالية:

➢ **المتغير الكمي المنفصل (المنقطع):**

$$V(x)=\frac{(x_1-\overline{X})^2 n_1+(x_2-\overline{X})^2 n_2+.......+(x_k-\overline{X})^2 n_k}{n_1+n_2+.......+n_k}$$

$$V(x)=\frac{\sum_{i=1}^{k}(x_i-\overline{X})^2 n_i}{\sum_{i=1}^{k} n_i}=\frac{\sum_{i=1}^{k} x_i^2 n_i}{\sum_{i=1}^{k} n_i}-\overline{X}^2$$

➢ **المتغير الكمي المتصل (المستمر):**

$$V(x)=\frac{(C_1-\overline{X})^2 n_1+(C_2-\overline{X})^2 n_2+.......+(C_k-\overline{X})^2 n_k}{n_1+n_2+.......+n_k}$$

$$V(x)=\frac{\sum_{i=1}^{k}(C_i-\overline{X})^2 n_i}{\sum_{i=1}^{k} n_i}=\frac{\sum C_i^2 n_i}{\sum_{i=1}^{k} n_i}-\overline{X}^2$$

أما الانحراف المعياري فهو جذر التباين $\sigma=\sqrt{V(x)}$

مثال (4-09): لتكن بيانات المثال السابق رقم (3-06) ، والتي وجدنا أن متوسطها الحسابي $\overline{X}=39{,}5$

الجدول رقم (4-07): توزيع مجموعة من الأفراد حسب دخلهم الشهري (الوحدة : 10^3دج)

فئات الدخل الشهري x_i	$[10-20[$	$[20-30[$	$[30-40[$	$[40-50[$	$[50-60[$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: إيجاد الانحراف المعياري للبيانات ؟

الحل: إعتمادا على جدول التوزيع التكراري يتم حساب التباين ومنه يتم إستخراج الانحراف المعياري وهذا بعد استخراج مراكز الفئات.

الجدول رقم (4-08): حساب الانحراف المتوسط اعتمادا على جدول التوزيع التكراري، حيث $\overline{X}=39{,}5$

الفئات X_i	n_i	C_i	$(C_i-\overline{X})$	$(C_i-\overline{X})^2 n_i$	$C_i^2 n_i$
$]20-10]$	5	15	24,5	3001,25	1125
$]30-20]$	8	25	14,5	1682	5000
$]40-30]$	12	35	4,5	243	14700
$]50-40]$	25	45	5,5	756,25	50625
$]60-50]$	10	55	15,5	2402,5	30250
المجموع	60	/	/	8085	101700

إذن قبل إستخراج الانحراف المعياري لابد من حساب التباين.

$$V(x)=\frac{\sum_{i=1}^{k}(C_i-\overline{X}\;)^2 n_i}{\sum_{i=1}^{k} n_i}=\frac{8085}{60}=134{,}75$$

$$V(x)=\frac{\sum C_i^{\,2}\; n_i}{\sum_{i=1}^{k} n_i}-\overline{X}^{\;2}=\frac{101700}{60}-(39{,}5)^2=134{,}75$$

إذن : $\sigma=\sqrt{V(x)}=\sqrt{134{,}75}=11{,}60$

IV-2-1-4-3. خواص الانحراف المعياري:

- يعتبر من أهم وأدق مقاييس التشتت وأكثرها استخداما؛
- يدخل في حسابه جميع مفردات القيم، وبالتالي يتأثر بالقيم المتطرفة؛
- لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة؛
- يقاس بوحدة القياس للمتغير الأصلي؛
- يمكن الاعتماد عليه أثناء المقارنة بين تشتت توزيعين إحصائيين من نفس النوعية ولهما نفس المتوسط الحسابي.

IV-2-2. مقاييس التشتت النسبية:

إن من بين مساوئ مقاييس التشتت المطلقة أنه يصعب إستخدامها في المقارنة بين سلسلتين أو أكثر إذا كانت وحدات قياسها مختلفة، وبالتالي كان البحث عن مقاييس تشتت أخرى لا تتأثر بوحدات القياس، هذه المقاييس هي مقاييس التشتت النسبية، والتي تتمثل في معامل الاختلاف النسبي ومعامل الاختلاف الربيعي.

IV-2-2-1. معامل الاختلاف النسبي (Coefficient of Variation):

هو من أفضل مقاييس التشتت النسبية وأكثرها إستخداما في حالة المقارنة بين تشتت سلسلتين أو أكثر ولها وحدات قياس مختلفة، وكذلك تختلف متوسطاتها الحسابية، يرمز له بالرمز $C.V$ ، ويعتمد في حسابه على المتوسط الحسابي للبيانات وإنحرافها المعياري، وتكون صيغنه الرياضية كما يلي:

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

مثال (4-10): من معطيات المثال رقم (4-09) وحله النموذجي وجدنا أن $\overline{X} = 39{,}5$ و $\sigma = 11{,}60$

المطلوب: أحسب معامل الاختلاف النسبي ؟ ماذا تلاحظ ؟

الحل: لدينا

$$C.V = \frac{\sigma}{\overline{X}} \times 100 = \frac{11{,}60}{39{,}5} \times 100 = 29{,}36$$

نلاحظ أن قيمة التشتت تقترب من 30% ، وبالتالي هناك تشتت قليل للبيانات، حيث أنه كلما إقترب من الصفر دل ذلك على تجانس البيانات، وكلما إقترب من 100% نقول أن البيانات مشتت كليا.

خواص معامل الاختلاف النسبي:

يتميز معامل الاختلاف النسبي بالخصائص التالية:

- لا يمكن قياسه للبيانات التي لها متوسط حسابي معدوم؛
- يقيس التشتت النسبي للبيانات دون وحدة قياسها؛
- يعتمد في حسابه على مقياس نزعة مركزية (المتوسط الحسابي) ومقياس تشتت (الانحراف المعياري)؛
- ليس له أهمية في المقارنة بين تشتت بيانات سلسلتين أو أكثر والتي لها نفس المتوسط الحسابي؛
- يستخدم خاصة لمقارنة تشتت بيانات سلسلتين أو أكثر من وحدات قياس مختلفة.
- لا يمكن قياسه من البيانات ذات الفئات المفتوحة لأنه لا يمكن حساب المتوسط الحسابي والانحراف المعياري.

IV-2-2-2. معامل الاختلاف (التباين) الربيعي (Coefficient Quartile Variation):

إن من مساوئ معامل الاختلاف النسبي أنه لا يمكن حسابه من البيانات ذات الفئات المفتوحة، وبالتالي وجد معامل آخر للتشتت النسبي يعالج هذه المشكلة هو معامل الاختلاف الربيعي الذي يرمز له بالرمز C.Q.V ، ويستخدم في حسابه كل الربيعيات، وبالتالي صيغته الرياضية هي :

$$C.Q.V = \frac{Q_3 - Q_1}{Q_2} \times 100$$

مثال (4-11): من معطيات المثال رقم (3-28) وحله النموذجي وجدنا أن $Q_1 = 31,66$ و $Q_2 = 42$ و $Q_3 = 48$

المطلوب: أحسب معامل الاختلاف الربيعي، ثم قارنه مع معامل ربيعي لبيانات أخرى قدره $C.Q.V_2 = 20\%$ ؟

الحل: لدينا

$$C.Q.V_1 = \frac{Q_3 - Q_1}{Q_2} \times 100 = \frac{48 - 31,66}{42} \times 100 = 38,9\%$$

نلاحظ أن $C.Q.V_2 < C.Q.V_1$ ، أي أن معامل الاخنلاف الربيعي المحسوب أكبر من المعطى، وبالتالي بيانات الأول أكثر تشتتا من الثاني.

المحور الخامس:

مقاييس الشكل

المحور الخامس : مقاييس الشكل

V-1. تمهيد: بالرغم من أن مقاييس التشتت تعطي لنا الشكل الانتشاري للبيانات سواءا انتشارها ما بين قيمها الكبرى والصغرى، أو من خلال تباعدها وتقاربها من متوسطاتها، إلا أنها لا تعطي لنا شكل الانتشار من حيث الجهة التي تتموقع فيها أغلبية البيانات بالمقارنة مع التموقع المنتظم (التوزيع الطبيعي)، وكذلك ابتعاد البيانات أو قربها من المحور الأفقي بالمقارنة مع الشكل المنتظم، ولتحقيق هذه المعرفة اكتشفت مقاييس أخرى لهذا الغرض سميت مقاييس الالتواء والتطاول (أو التفلطح).

إن مقاييس الالتواء والتطاول (أو التفلطح) هي مقاييس تحدد شكل إنتشار البيانات بالمقارنة مع الشكل الانتشاري الطبيعي الذي يتساوى فيه المتوسط الحسابي مع الوسيط والمنوال سواءا من حيث الميلان أو الالتواء، وكذلك من حيث الارتفاع (التطاول والتفلطح)، ونظرا لكون أن هذه المقاييس يعتمد في حسابها على العزوم البسيطة أو المركزية فسنقوم أولا بعرض طرق حساب هذه العزوم.

V-2. العزوم :

يوجد نوعان من العزوم، عزوم تدور حول المبدأ (الصفر) والتي تسمى بالعزوم البسيطة، و يرمز لها بـ m_r وأخرى تدور حول المتوسط الحسابي والتي تسمى بالعزوم المركزية، و يرمز لها بـ M_r ، أما رتبة (رقم) العزم (r) فيتحدد وفقها القوة (الأس) التي ترفع إليها القيم أو انحرافاتها عن المتوسط الحسابي.

V-2-1. حساب العزوم للبيانات الأولية (غير المبوبة) :

إذا كانت قيم الظاهرة المدروسة $x_1 ; x_2 ; ... ; x_n$ والتي عددها $\mathbf{n}$ فإن االعزوم يتم حسابها وفق الصيغ الموافقة لها.

- **العزوم البسيطة**: تحسب وفق العلاقة التالية

$$m_r = \frac{\sum x_i^r}{n}$$

- **العزوم المركزية**: تحسب وفق العلاقة التالية

$$M_r = \frac{\sum \left(x_i - \overline{X}\right)^r}{n}$$

حيث : r يمثل رقم (رتبة) العزم.

مثال (5-01): لتكن بيانات المثال السابق رقم (3-01) ، والتي وجدنا أن متوسطها الحسابي $\overline{X} = 7{,}5$

2 ، 4 ، 8 ، 9 ، 10 ، 8 ، 10 ، 9 ، 10 ، 7 ، 5 ، 8

المطلوب: أحسب العزم البسيط الأول والعزم المركزي الثاني ؟ ماذا تلاحظ ؟

الحل: لدينا

$$m_1 = \frac{\sum x_i^1}{n} = \frac{2+4+8+9+10+8+10+9+10+7+5+8}{12} = \frac{90}{12} = 7{,}5$$

$$M_2 = \frac{\sum \left(x_i - \overline{X}\right)^2}{n}$$

$$M_2 = \frac{(2-7{,}5)^2 + (4-7{,}5)^2 + (8-7{,}5)^2 + (9-7{,}5)^2 + (10-7{,}5)^2 + (8-7{,}5)^2 + (10-7{,}5)^2 + (9-7{,}5)^2 + (10-7{,}5)^2 + (7-7{,}5)^2 + (5-7{,}5)^2 + (8-7{,}5)^2}{12}$$

$$M_2 = \frac{30{,}25+12{,}25+0{,}25+2.25+6{,}25+0{,}25+6{,}25+2{,}25+6{,}25+0{,}25+6{,}25+0{,}25}{12}$$

$$M_2 = \frac{73}{12} = 6{,}08$$

نلاحظ أن : $\overline{X} = m_1$

كذلك وجدنا سابقا من نفس البيانات أن : $V(x) = \frac{73}{12} = 6{,}08$ ، وبالتالي : $M_2 = V(x)$

V-2-2. حساب العزوم للبيانات المبوبة:

إذا كانت القيم النقطية للمتغير الكمي المنفصل هي : $x_1 ; x_2 ; .... ; x_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، أو إذا كانت مراكز الفئات للمتغير الكمي المستمر هي : $C_1 ; C_2 ; .... ; C_k$ ، وكانت $n_1 ; n_2 ; .... ; n_k$ تمثل التكرارات المقابلة لها، وكان متوسطها الحسابي $\overline{X}$ فإن االعزوم يتم حسابها وفق الصيغ الموافقة لها.

- **العزوم البسيطة**: تحسب وفق العلاقة التالية $m_r = \frac{\sum x_i^r n_i}{\sum n_i}$ أو $m_r = \frac{\sum C_i^r n_i}{\sum n_i}$

- **العزوم المركزية**: تحسب وفق العلاقة التالية $M_r = \frac{\sum (x_i - \overline{X})^r n_i}{\sum n_i}$ أو $M_r = \frac{\sum (C_i - \overline{X})^r n_i}{\sum n_i}$

حيث : r يمثل رقم (رتبة) العزم.

مثال (5-01): لتكن بيانات المثال السابق رقم (3-06) ، والتي وجدنا أن متوسطها الحسابي $\overline{X} = 39{,}5$

الجدول رقم (5-01): توزيع مجموعة من الأفراد حسب دخلهم الشهري.

فئات الدخل الشهري x_i	$]10-20]$	$]20-30]$	$]30-40]$	$]40-50]$	$]50-60]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

المطلوب: أحسب العزم البسيط الثاني والعزمين المركزيين الثالث والرابع؟

الحل: إعتمادا على جدول التوزيع التكراري يتم حساب العزوم الابتدائية والمركزية وهذا بعد استخراج مراكز الفئات.

الجدول رقم (5-02): حساب العزوم اعتمادا على جدول التوزيع التكراري.

الفئات	n_i	C_i	C_i^2	$C_i^2 n_i$	$(C_i - \overline{X})$	$(C_i - \overline{X})^3 n_i$	$(C_i - \overline{X})^4 n_i$
$]10-20]$	5	15	225	1125	-24,5	-73530,625	1801500,313
$]20-30]$	8	25	625	5000	-14,5	-24389	353640,5
$]30-40]$	12	35	1225	14700	-4,5	-1093,5	4920,75
$]40-50]$	25	45	2025	50625	5,5	4159,375	22876,5625
$]50-60]$	10	55	3025	30250	15,5	37238,75	577200,625
المجموع	60	/	/	101700	/	-57615	2760138,7505

- العزم البسيط الثاني:

$$m_2 = \frac{\sum C_i^2 n_i}{\sum n_i} = \frac{101700}{60} = 1695$$

- العزم المركزي الثالث:

$$M_3 = \frac{\sum \left(C_i - \overline{X}\right)^3 n_i}{\sum n_i} = \frac{-57615}{60} = -960{,}25$$

- العزم المركزي الرابع:

$$M_4 = \frac{\sum \left(C_i - \overline{X}\right)^4 n_i}{\sum n_i} = \frac{2760138{,}7505}{60} = 46002{,}312$$

V–3. الالتواء (Skewness):

إن الالتواء يعني عدم انتظام البيانات، فقد تكون أغلبيتها تتمركز في بداية المنحنى (أكبر تكرارات القيم النقطية أو الفئات تكون في البداية)، أو في نهايته (أكبر تكرارات القيم النقطية أو الفئات تكون في النهاية) بالمقارنة بالتمثيل المتماثل التي تكون بياناته تتمركز في الوسط.

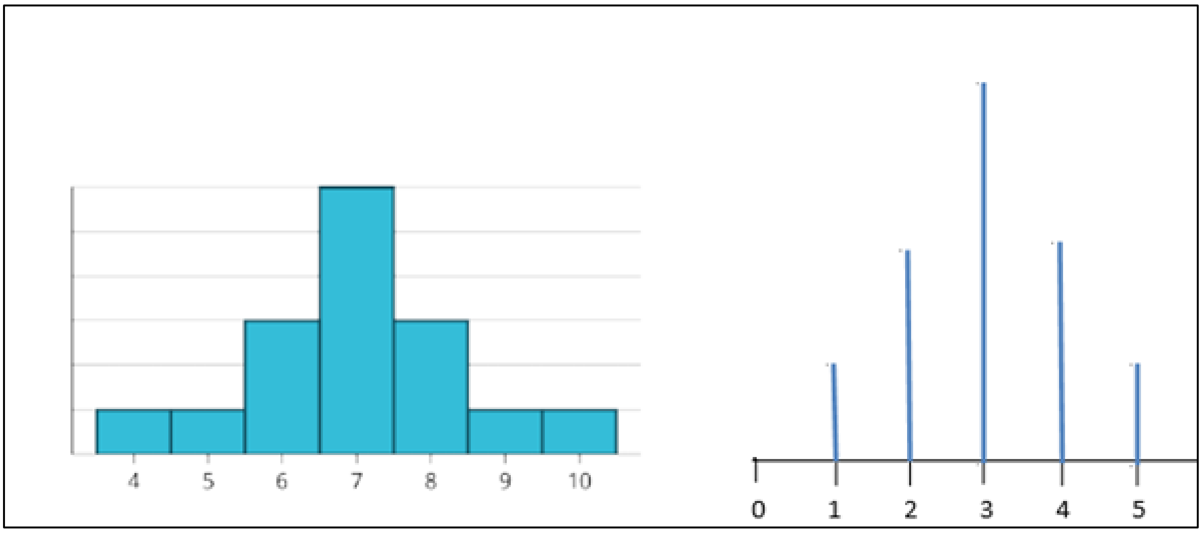
كما يمكن تعريفه على أنه " درجة اللا تماثل، أو الابتعاد عن التناظر، لتوزيع معين. إذا كانت منحنى التكرار لتوزيع ما له ذيل أطول على يمين الحد الأقصى المركزي مقارنة باليسار، يُقال إن التوزيع مائل إلى اليمين، أو له إلتواء موجب. وإذا كان العكس صحيحًا، يُقال إن التوزيع مائل إلى اليسار، أو له إلتواء سلبي[1].

V–3–1. أشكال الالتواء:

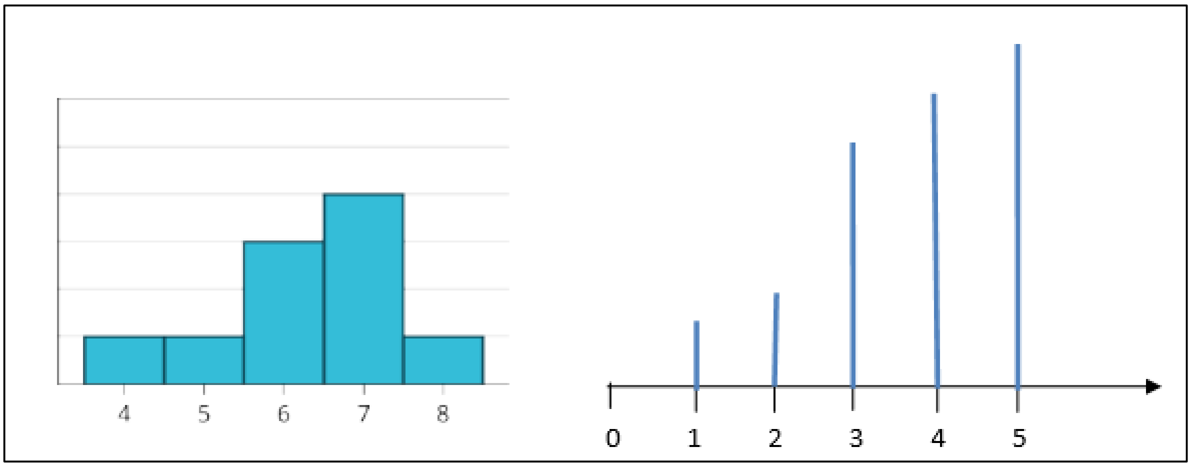
يوجد شكلين للالتواء ، فالالتواء قد يكون في جهة اليسار، وقد يكون في جهة اليمين، بالمقارنة بالتمثيل المتماثل الذي يأخذ الشكل التالي:

[1] - MURRAY R. SPIEGEL and LARRY J. STEPHENS ; **Theory and Problems of STATISTICS ;** Schaum's Outline Series ; Fourth Edition ; 2008 ; P 125.

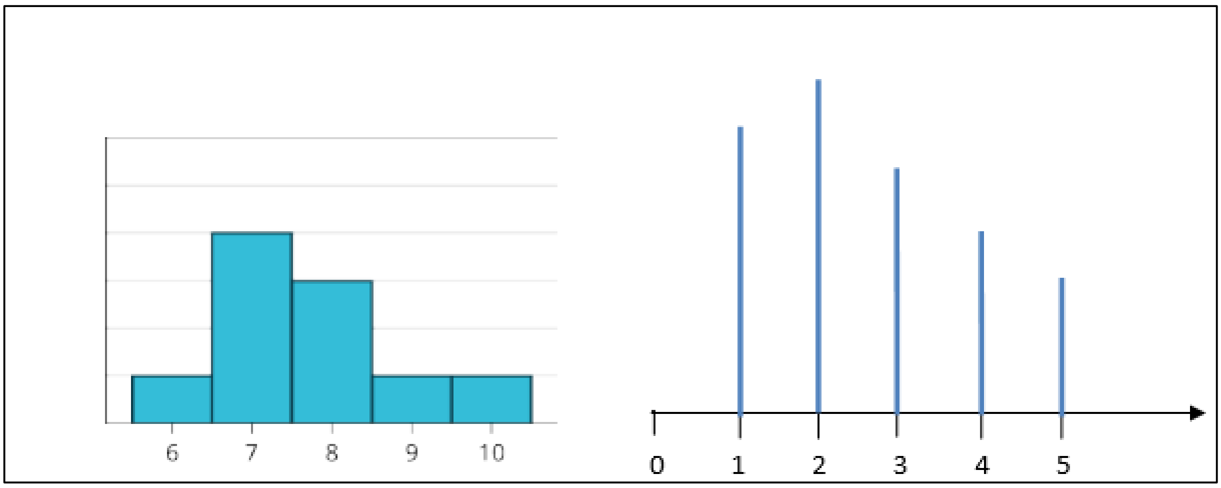
الشكل رقم (5–01): الأشكال البيانية الممكنة للتمثيل المتماثل (المتناظر)



الشكل رقم (5–02): الأشكال البيانية الممكنة للتمثيل الملتوي نحو اليسار



الشكل رقم (5–03): الأشكال البيانية الممكنة للتمثيل الملتوي نحو اليمين



V-3-2. قياس الالتواء :

يقاس الالتواء بعدة مقاييس أبرزها معاملات بيرسون، معامل فيشر ومعامل يول.

V-3-2-1. معاملات بيرسون للالتواء:

توجد ثلاثة أنواع من معاملات بيرسون يمكن من خلالها يتم تحديد شكل الالتواء.

V-3-2-1-1. معامل بيرسون الأول P_1: يعتمد هذا المعامل على ثلاثة مقاييس هي المتوسط الحسابي، المنوال والانحراف المعياري وفق العلاقة التالية:

$$P_1 = \frac{\overline{X} - M_0}{\sigma}$$

والقرار يكون حسب قيمة المعامل :

- إذا كان: $P_1 = 0$ ، أي $\overline{X} = M_o$ فإن منحنى التوزيع يكون متماثل.
- إذا كان: $P_1 > 0$ فإن منحنى التوزيع يكون مائل لليمين.
- إذا كان: $P_1 < 0$ فإن منحنى التوزيع يكون مائل لليسار.

V-3-2-1-2. معامل بيرسون الثاني P_2: يعتمد هذا المعامل كذلك على ثلاثة مقاييس هي المتوسط الحسابي، الوسيط والانحراف المعياري وفق العلاقة التالية:

$$P_2 = \frac{3 \times \left(\overline{X} - M_e\right)}{\sigma}$$

والقرار يكون حسب قيمة المعامل :

- إذا كان: $P_2 = 0$ ، أي $\overline{X} = M_e$ فإن منحنى التوزيع يكون متماثل.
- إذا كان: $P_2 > 0$ فإن منحنى التوزيع يكون مائل لليمين.
- إذا كان: $P_2 < 0$ فإن منحنى التوزيع يكون مائل لليسار.

ملاحظة: يوجد معامل بيرسون آخر يعتمد على العزوم وهو معامل بيرسون العزمي $\beta_1 = \frac{M_3^2}{M_2^3}$ ولكنه غير مجد لأنه دوما موجب وبالتالي لا يعطي لنا شكل التوزيع، ويستخدم في معرفة أن التوزيع منتظم أو غير منتظم فقط.

مثال (5-02): لتكن البيانات السابقة الخاصة بمجموعة من الأفراد تم دراسة دخلهم الشهري.

الجدول رقم (5-03): توزيع مجموعة من الأفراد حسب دخلهم الشهري.

فئات الدخل الشهري x_i	$]10-20]$	$]20-30]$	$]30-40]$	$]40-50]$	$]50-60]$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

وقد وجدنا أن $\overline{X} = 39{,}5$ و $M_e = 42$ و $M_0 = 44{,}64$ و $\sigma = 11{,}60$

المطلوب: ما هو شكل التوزيع حسب معاملي بيرسون للالتواء؟

- معامل بيرسون الأول P_1 :

$$P_1 = \frac{\overline{X} - M_0}{\sigma} = \frac{39,5 - 44,64}{11,60} = -0,44$$

- معامل بيرسون الثاني P_2 :

$$P_2 = \frac{3 \times \left(\overline{X} - M_e\right)}{\sigma} = \frac{3(39,5 - 42)}{11,60} = -0,64$$

نلاحظ أن معاملي بيرسون الأول والثاني كلاهما سالب وبالتالي منحنى التوزيع يلتوي نحو اليسار.

V-3-2-2. معامل فيشر للالتواء (F_1) : يعتمد هذا المعامل على العزم المركزي الثالث والانحراف المعياري، ويعطى بالصيغة الرياضية التالية :

$$F_1 = \frac{M_3}{\sigma^3}$$

حيث: M_3 هو العزم المركزي الثالث و σ هو الانحراف المعياري.

وشكل التوزيع متوقف على القرار التالي :

- إذا كان: $F_1 = 0$ فإن منحنى التوزيع يكون متماثل، لأن العزم المركزي الثالث في هذه الحالة معدوم ($M_3 = 0$)
- إذا كان: $F_1 > 0$ فإن منحنى التوزيع يكون مائل لليمين.
- إذا كان: $F_1 < 0$ فإن منحنى التوزيع يكون مائل لليسار.

مثال (5-03): من معطيات المثال السابق

الجدول رقم (5-04): توزيع مجموعة من الأفراد حسب دخلهم الشهري.

فئات الدخل الشهري x_i	$[10-20[$	$[20-30[$	$[30-40[$	$[40-50[$	$[50-60[$	المجموع
عدد الأفراد n_i	5	8	12	25	10	60

وجدنا أن $M_3 = 960,25$ و $\sigma = 11,60$

المطلوب: ما هو شكل التوزيع حسب معامل فيشر للالتواء؟

الحل:

$$F_1 = \frac{M_3}{\sigma^3} = \frac{-960,25}{11,60^3} = -0,61$$

نلاحظ أن $F_1 < 0$ وبالتالي منحنى التوزيع يميل (يلتوي) نحو اليسار.

V-3-2-3. معامل يول للالتواء (C_Y) : يستخدم هذا المعامل في حالة جداول التوزيع التكرارية المفتوحة، والتي من خلالها لا يمكن حساب باقي معاملات الالتواء، وهو يعتمد على كل الربيعيات، ويعطى بالصيغة الرياضية التالية :

$$C_Y = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

وشكل التوزيع متوقف على القرار التالي :

- إذا كان: $C_Y = 0$ فإن منحنى التوزيع يكون متماثل.
- إذا كان: $C_Y > 0$ فإن منحنى التوزيع يميل نحو اليمين.
- إذا كان: $C_Y < 0$ فإن منحنى التوزيع يميل نحو اليسار.

مثال (5–04): من معطيات حل المثال رقم (3–26) وجدنا قيم الربيعيات.

$$Q_3 = 48 \quad و \quad Q_2 = 42 \quad و \quad Q_1 = 31{,}66$$

المطلوب: ما هو شكل التوزيع حسب معامل يول للالتواء؟

الحل:

$$C_y = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{(48-42)-(42-31{,}66)}{48-31{,}66} = \frac{-4{,}34}{16{,}34} = -0{,}26$$

نلاحظ أنّ : $C_Y < 0$ وبالتالي منحنى التوزيع يميل نحو اليسار.

ملاحظة : لقد وجدنا من خلال نفس معطيات المثال أن كل معاملات الالتواء حددت شكل التوزيع أنه يميل نحو اليسار.

V–4. التفلطح والتطاول (Kurtosis):

التفلطح والتطاول هو قياس درجة علو قمة التوزيع بالنسبة للتوزيع الطبيعي، أي يقصد به مدى اتساع أو ضعف قمة منحنى التوزيع، فكلما كان الشكل أكثر ارتفاعا من الشكل الطبيعي نقول أن الشكل مدبب، أما إذا كان أقل ارتفاع من الشكل الطبيعي فنقول عنه أنه مفلطح.

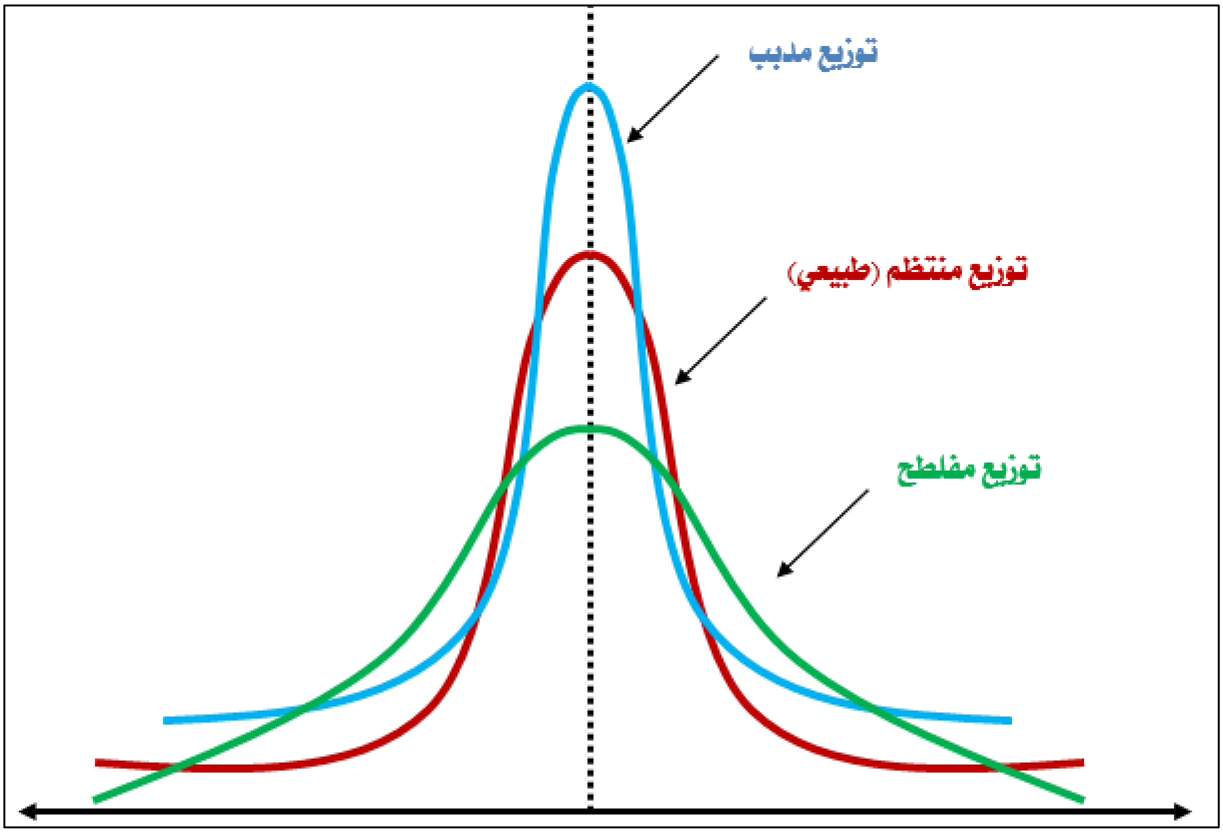
كما يمكن تعريف التفلطح والتطاول على أنه " درجة حدَّة القمّة في التوزيع، وعادةً ما يتم قياسه بالنسبة للتوزيع الطبيعي. يُسمى التوزيع الذي له قمة نسبيًا مرتفعة بالتوزيع الحاد القمة (مدبب)، وعادة ما تكون له قاعدة أقل إتساع، بينما يُسمى التوزيع الذي له قمة نسبيا منخفضة بالتوزيع المسطح القمة، وعادة تكون له قاعدة متسعة بالمقارنة مع التوزيع الطبيعي، الذي ليس حاد القمة جدًا ولا مسطح القمة جدًا.[2]

V–4–1. أشكال التفلطح والتطاول:

إنطلاقا من تعريف التفلطح والتطاول يمكننا وضع تمثيلات بيانية توضح أشكال التوزيعات الممكنة

[2] - MURRAY R. SPIEGEL and LARRY J. STEPHENS ; Op – Cit ; P

الشكل رقم (5–04): أنواع التمثيلات الممكنة (التفلطح والتطاول) للتوزيعات التكرارية .



V–4–2. قياس التفلطح والتطاول:

يتم قياس التفلطح والتطاول بأحد المقاييس التالية:

V–4–2–1. معامل بيرسون للتفلطح β_2 : يعتمد هذا المعامل على العزمين المركزيين الرابع والثاني، وصيغته الرياضية تعطى كما يلي:

$$\beta_2 = \frac{M_4}{M_2^2} = \frac{M_4}{[V(x)]^2}$$

وشكل التوزيع متوقف على القرار التالي :

- إذا كان: $B_2 = 3$ فإن منحنى التوزيع يكون متماثل (طبيعي).
- إذا كان: $B_2 > 3$ فإن منحنى التوزيع مدبب (متطاول).
- إذا كان: $B_2 < 3$ فإن منحنى التوزيع مفلطح.

V–4–2–2. معامل فيشر للتفلطح F_2: هو معامل يعتمد كذلك على العزمين المركزيين الرابع والثاني، وهو عبارة عن معامل بيرسون مطروحا منه القيمة 3 ، أما صيغته الرياضية فهي كالتالي:

$$F_2 = \frac{M_4}{M_2^2} - 3 = \beta_2 - 3$$

وشكل التوزيع متوقف على القرار التالي :

- إذا كان: $F_2 = 0$ فإن منحنى التوزيع يكون متماثل (طبيعي)، لأن $B_2 = 3$.
- إذا كان: $F_2 > 0$ فإن منحنى التوزيع مدبب (متطاول).
- إذا كان: $F_2 < 0$ فإن منحنى التوزيع مفلطح.

مثال (5–05): لتكن بيانات المثال (5–01) والتي وجدنا فيها أن : $M_4 = 46002{,}312$ ووجدنا سابقا لنفس المعطيات أن $V(x) = 134{,}75$.

المطلوب: هل التوزيع مفلطح أم مدبب وفق معاملي بيرسون وفيشر؟

الحل:

$$\beta_2 = \frac{M_4}{M_2^2} = \frac{M_4}{[V(x)]^2} = \frac{46002{,}312}{(134{,}75)^2} = 2{,}53$$

نلاحظ أن معامل بيرسون للتفلطح أقل من 3 ، $(B_2 < 3)$ وبالتالي فإن منحنى التوزيع مفلطح.

$$F_2 = \frac{M_4}{M_2^2} - 3 = \beta_2 - 3 = 2{,}53 - 3 = -0.47$$

كذلك نلاحظ أن معامل فيشر للتفلطح سالب ، $(F_2 < 0)$ وبالتالي منحنى التوزيع مفلطح.

V–4–2–2. معامل التفلطح المئوي (β): هو معامل يعتمد عليه في معرفة شكل التوزيع هل هو مفلطح أو مدبب في حالة متغير كمي مستمر وجدول توزيعه التكراري مفتوح، وهو يقاس إعتمادا على المدى الربيعي والمئويين العاشر والتسعون وفق الصيغة الرياضية التالية:

$$\beta = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} = \frac{I_Q}{2(P_{90} - P_{10})}$$

وشكل التوزيع متوقف على القرار التالي:

- إذا كان: $\beta = 0{,}263$ فإن منحنى التوزيع يكون متماثل (طبيعي).
- إذا كان: $\beta > 0{,}263$ فإن منحنى التوزيع مدبب (متطاول).
- إذا كان: $\beta < 0{,}263$ فإن منحنى التوزيع مفلطح.

مثال (5–06): لتكن بيانات جدول التوزيع التكراري المفتوح الموالي

الجدول رقم (5–05): بيانات جدول توزيع تكراري مفتوح

الفئات	أقل من 20	$[20-30[$	$[30-40[$	$[40-50[$	$[50-60[$	المجموع
n_i	5	8	12	25	10	60

المطلوب: ما هو شكل التوزيع من خلال التفلطح أو التطاول ؟

الحل: بما أن جدول التوزيع التكراري مفتوح فلا يمكن معرفة شكل التوزيع (مفلطح أو مدبب) عن طريق معاملي بيرسون أو فيشر، بل يستخدم معامل التفلطح المئوي (β)

الجدول رقم (5-06): **إستخدام** جدول التوزيع التكراري في حساب الربيعيين الأول والثالث والمئويين العاشر والتسعون.

الفئات	n_i	$n_i \uparrow$
أقل من 20	5	5
$]30-20]$	8	13
$]40-30]$	12	25
$]50-40]$	25	50
$]60-50]$	10	60
المجموع	60	/

الربيع الأول:

رتبة الربيع الأول (Q_1): $\frac{iN}{4}=\frac{1(60)}{4}=15$، إذن فئة الربيع الأول هي: $]40-30]$

وقيمة الربيع الأول هي:

$$Q_1=L_1+\frac{\frac{iN}{4}-N_0}{n_{Q_1}}\times k=30+\frac{\frac{1(60)}{4}-13}{12}\times 10=30+\frac{15-13}{12}\times 10=31{,}66$$

الربيع الثالث:

رتبة الربيع الثالث (Q_3): $\frac{iN}{4}=\frac{3(60)}{4}=45$، إذن فئة الربيع الأول هي: $]50-40]$

وقيمة الربيع الثالث هي:

$$Q_3=L_1+\frac{\frac{iN}{4}-N_0}{n_{Q_3}}\times k=40+\frac{\frac{3(60)}{4}-25}{25}\times 10=40+\frac{45-25}{25}\times 10=48$$

المئوي العاشر:

رتبة المئوي العاشر (P_{10}): $\frac{iN}{100}=\frac{10(60)}{100}=6$، إذن فئة المئوي العاشر هي: $]30-20]$

وقيمة المئوي العاشر هي:

$$P_{10}=L_1+\frac{\frac{iN}{100}-N_0}{n_{P_{10}}}\times k=20+\frac{\frac{10(60)}{100}-5}{8}\times 10=20+\frac{6-5}{8}\times 10=21{,}25$$

المئوي التسعون:

رتبة المئوي التسعون (P_{90}): $\frac{iN}{100}=\frac{90(60)}{100}=54$، إذن فئة المئوي التسعون هي: $]60-50]$

وقيمة المئوي التسعون هي:

$$P_{90}=L_1+\frac{\frac{iN}{100}-N_0}{n_{P_{90}}}\times k=50+\frac{\frac{90(60)}{100}-50}{10}\times 10=50+\frac{54-50}{10}\times 10=54$$

إذن معامل التفرطح المئوي هو:

$$\beta = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} = \frac{48 - 31{,}66}{2(54 - 21{,}25)} = \frac{16{,}34}{65{,}5} = 0{,}249$$

وجدنا أن $\beta = 0{,}249 < 0{,}263$ فن منحنى التوزيع مفلطح.

المحور السادس:

مقاييس التمركز

(منحنى لورنز ومعامل جيني)

VI-1. تمهيد:

قام أب عائلة مكونة من 7 أولاد بتوزيع مبلغ مالي قدره 800 دينار من أجل احتياجاتهم الخاصة ، فتم التوزيع وفق ترتيبهم عمريا من أكبر سن إلى أقلهم سنا وفق التسلسل الموالي : أخذ أحمد 200 دينار ، وأخذت لياء 150 دينار ثم أخذت تقوى 120 دينار ، وأعطي سامي 100 دينار ، ثم أخذ رضا و أسماء 90 دينار و 80 دينار على التوالي ، وأخذ الطفل الأقل سنا حمزة 60 دينار.

إن السؤال المطروح هو : هل كان الأب عادلا في توزيع المبلغ المالي على أولاده ؟

من الوهلة الأولى يظهر أن الأب لم يقم بالتوزيع العادل أو المتساوي لمبلغه المالي على أولاده ، لكن عندما نطلع على حاجيات الأولاد من الأكبر إلى الأصغر يمكن القول أنه أظهر نوع من المساواة ، ليبقى التساؤل العام : كيف يمكن قياس المساواة ؟

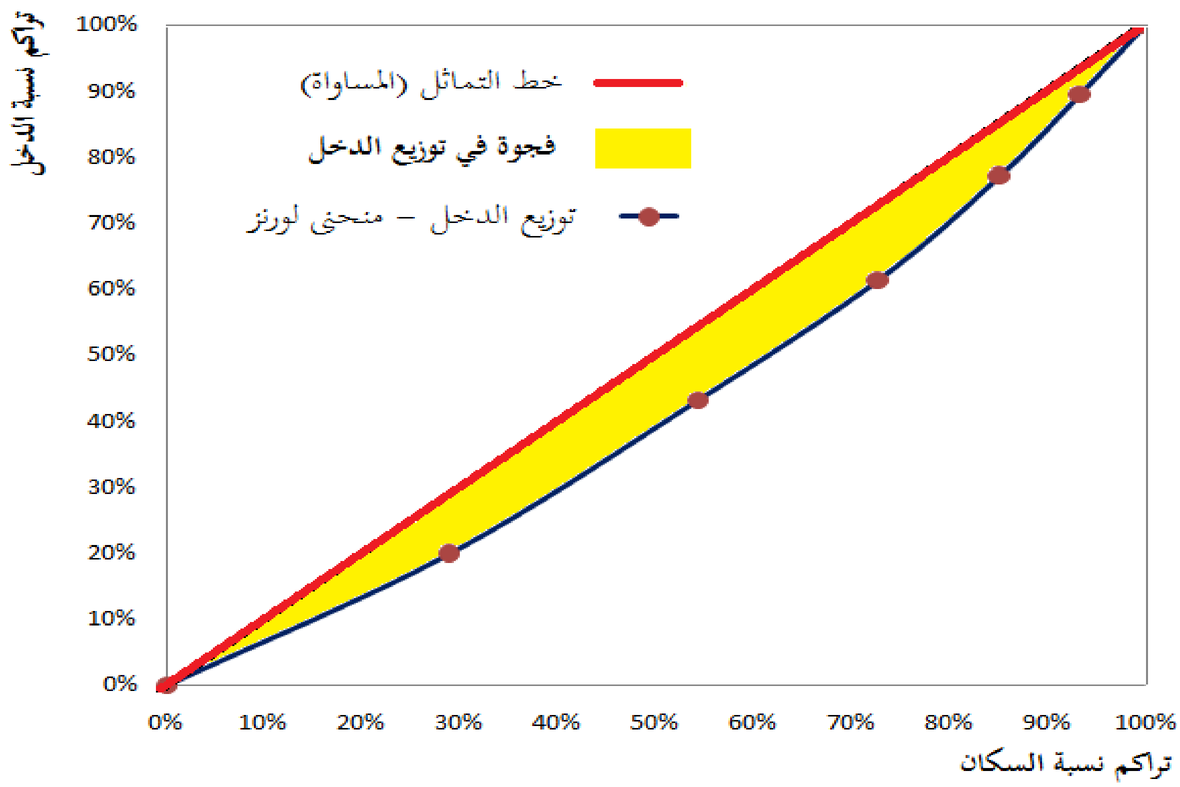
غالبا ما يقارن الناس بين أوضاعهم المالية الشخصية وأوضاع جيرانهم أو زملائهم في العمل أو أصدقائهم على أساس مساكنهم التي يقطنوها أو ما لديهم من مقتنيات، وعادة ما يستخدم الاقتصاديون المسوح الأسرية لقياس مدى التفاوت في الدخول، فتُجرى مقابلات مع طائفة كبيرة من الأسر لتحديد مصادر دخولها المختلفة (النقدية والنوعية) وأنماطها الاستهلاكية، ثم يُقسَّم مجموع دخل الأسرة بعد خصم الضرائب المباشرة المدفوعة (أو إجمالي استهلاك الأسرة) على عدد الأفراد المقيمين ضمن الأسرة الواحدة ثم يصنف جميع المشمولين في المسح في مراتب، من الأفقر إلى الأغنى، وفقا للدخل الأسري للفرد ، إن هذا الأسلوب الرياضي يتيح لنا حساب توزيع الدخل والذي قد يكون عادلا أو غير عادل.

إن التوزيع العادل للدخل أو أي ظاهرة ما يقاس بمقياسين هامين يسميان بمقاييس التمركز، أي تمركز توزيع الفعلي للظاهرة عن خط التماثل (التعادل)، وهاذين المقياسين هما منحنى التمركز والذي يعرف بمنحنى لورنز (LORENZ Curve) ومؤشر التمركز الذي يعرف بمعامل جيني (GINI Coefficient).

VI-2. منحنى التمركز:

VI-2-1. تعريف منحنى التمركز: هو منحنى معروف بمحنى لورنز (LORENZ Curve)، وهو أهم طريقة لتمثيل توزيع الدخل بيانيا، يستخدم لغرض تمثيل التفاوت في توزيع الدخل أو الإنفاق أو متغيرات أخرى بشكل بياني ، فيرسم هذا المنحنى ضمن مربع طول ضلعه يمثل 100% ، فالمحور الأفقي له يمثل المتجمع الصاعد للنسب المئوية لعدد الأفراد والمحور العمودي يمثل المتجمع الصاعد للنسب المئوية للدخل ، ويقسم هذا المنحنى بخط 45° الذي يسمى بخط المساواة أو خط الأمثلية ، وخط القطر الواصل ما بين الزاوية أسفل المربع من اليسار والزاوية أعلى المربع من اليمين كما يمثله التمثيل الموالي :

الشكل رقم (6-01) : التمثيل البياني لمنحنى التمركز (منحنى لورنز)

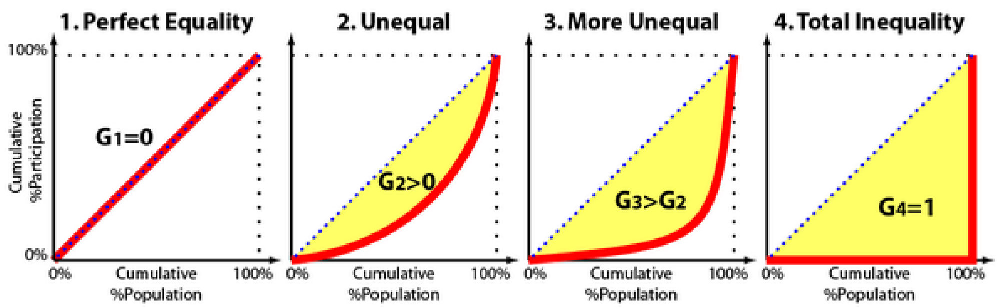


إن خط المساواة التام تتساوى فيه النسب المئوية لعدد الأفراد مع النسب المئوية للدخل ، وعادة ما يكون توزيع نسب الدخل أقل من توزيع نسب عدد الأفراد (السكان) وهو ما أدى إلى تواجد خط توزيع الدخل الحقيقي على يمين خط التماثل .

VI-2-2. بعض أشكال منحنى التمركز:

توجد عدة أشكال لمنحنى التمركز أبرزها :

الشكل رقم (6-02) : التمثيلات البيانية الممكنة لمنحنى التمركز (لورنز)



- التمثيل البياني الذي لا تكون فيه فجوة أي يتطابق منحنى لورنز مع خط التماثل وهو يعني وجود مساواة تامة في توزيع الثروة أو الدخل أو........ ، وهو معبر بالشكل 1 (Perfect Equality).
- التمثيل البياني الذي تكون فيه فجوة قليلة ما بين منحنى لورنز وخط التماثل، والذي يعني عدم وجود مساواة وإختلاف قليل في توزيع الثروة أو الدخل أو ....، وهو معبر بالشكل 2 (Unequal).

- التمثيل البياني الذي تكون فيه فجوة أكثر ما بين منحنى لورنز وخط التماثل، والذي يعني عدم وجود مساواة وإختلاف كبير في توزيع الثروة أو الدخل أو ....، وهو معبر بالشكل 3 (More Unequal).
- التمثيل البياني الذي تكون فيه فجوة تامة ما بين منحنى لورنز وخط التماثل، والذي يعني عدم وجود مساواة مطلقا ، وقد يعبر كذلك بأن أحد الأفراد أخذ كل الثروة أو الدخل أو .....دون غيره، وهو معبر بالشكل 4 (Total Inequality

VI-2-3. خواص منحنى التمركز (لورنز) : يتميز منحنى لورنز بالخواص التالية :

- يستخدم في معرفة درجة تمركز البيانات وتحديد الشكل العام لاتجاه توزيع الظواهر والتعرف على مدى عدالة التوزيع لظاهرة ما ؛
- يستخدم عند المقارنة بين التوزيع الفعلي للظاهرة وتوزيعها المثالي الذي ينطبق على خط التعادل (التماثل) ؛
- يعتمد على وجود متغيرين بوحدات قياس مختلفة:

أ- المتغير المستقل X_i (السكان ، الأفراد ، الملاك ، العمال ...) ؛

ب- المتغير التابع Y_i (المساحة ، الدخل ، الثروة ، الأراضي الزراعية ...) ؛

- كلما اقترب منحنى التوزيع الفعلي من خط التماثل دل ذلك على قرب توزيع الظاهرة من المثالية ؛
- يبدأ منحنى لورنز دائما من (0 ، 0) وينتهي عند (100% ، 100%) عند استخدام التكرار النسبي المئوي الصاعد (f_i %⇧) وعند (1 ، 1) عند استخدام التكرار النسبي الصاعد (f_i ⇧) لكلا المتغيرين.
- لا يمكن أن يرتفع منحنى لورنز فوق خط المساواة التامة.

VI-2-4. خطوات بناء منحنى التمركز (لورنز) :

من أجل رسم منحنى لورنز لابد من حساب المتجمع الصاعد للنسب المئوية لعدد الأفراد و المتجمع الصاعد للنسب المئوية للدخل أو الثروة أو........، وقد نصادف نوعين من المتغيرة قد تكون منفصلة (نقطية) أو مستمرة (فئات)

VI-2-4-1. خطوات بناء منحنى التمركز (لورنز) في حالة قيم نقطية (منفصلة) :

لتوضيح خطوات رسم منحنى لورنز في حالة وجود المتغيرة المنفصلة (نقطية) نعرض المثال الموالي.

مثال رقم (6-01) : تم توزيع مبلغ مالي على مجموعة من الأشخاص ، فاخذ نوح مبلغ 100 وحدة نقدية ، وأخذ عيسى 25 وحدة نقدية ، ثم سمير 25 وحدة نقدية ، ثم آدم 50 وحدة نقدية ، وإسلام أخذ 25 وحدة نقدية ، ثم محمد 25 وحدة نقدية ، ثم أخذ علي 50 وحدة نقدية ، وأخذ إسحاق 50 وحدة نقدية ، ثم عبد الرحمان أخذ 100 وحدة نقدية ، ثم إستلم عمر 150 وحدة نقدية .

المطلوب : هل تم توزيع المبلغ المالي توزيعا عادلا ؟

الحل: من أجل معرفة هل أن التوزيع عادلا أم لا نستخدم منحنى لورنز ، ويكون ذلك من خلال حساب التجمع التراكمي لنسب الأفراد والتجمع التراكمي لنسب الثروة ، لكن في البداية لابد من ترتيب الأشخاص حسب المبلغ المقدم لهم من أصغر مبلغ إلى الأكبر، وذلك موضح في الجدول الموالي :

الجدول رقم (6-01): الجدول التوزيع التكراري المستخدم في رسم منحنى لورنز

الفرد X_i	المبلغ (الثروة) Y_i	نسبة الفرد	نسبة المبلغ	التوزيع التراكمي لنسبة الأفراد P_i	التوزيع التراكمي لنسبة الثروة Q_i
عيسى	25	% 10	% 4.17	% 10	% 4.17
سمير	25	% 10	% 4.17	% 20	% 8.34
إسلام	25	% 10	% 4.17	% 30	% 12.48
محمد	25	% 10	% 4.17	% 40	% 16.65
آدم	50	% 10	% 8.33	% 50	% 24.98
علي	50	% 10	% 8.33	% 60	% 33.31
إسحاق	50	% 10	% 8.33	% 70	% 41.64
نوح	100	% 10	% 16.66	% 80	% 58.34
عبد الرحمان	100	% 10	% 16.66	% 90	% 75
عمر	150	% 10	% 25	% 100	% 100
المجموع	600	% 100	% 100	/	/

-التمثيل البياني لمنحنى لورنز :

الشكل رقم (6-02) : الرسم البياني لمنحنى التمركز (لورنز)

من خلال المنحنى يظهر هناك تفاوت في توزيع المبلغ (الثروة) لأن منحنى توزيع المبلغ يبتعد عن خط المساواة ، فلو كان التوزيع يتطابق مع خط التماثل لقلنا أن توزيع المبلغ على الأشخاص متساوي ، ولو وجدنا التوزيع يصل إلى 100 % عند منحنى التوزيع التراكمي لنسب الأفراد لقلنا أن شخص واحد إستحوذ على المبلغ المعطى ، لكن ما يميز هذا المنحنى هو أنه لا يعطينا قيمة للتفاوت.

VI-2-4-2. خطوات بناء منحنى التمركز (لورنز) في حالة فئات (مستمر) :

إن خطوات رسم منحنى لورنز في حالة وجود فئات هي مماثلة لسابقتها مع الأخذ بمراكز الفئات كأرقام نقطية كما هو مبين في المثال الموالي.

مثال رقم (6-02) : إذا تم تقسيم عدد عمال مؤسسة حسب أجورهم الشهرية كما هو مبين في الجدول اللاحق، وطلب رسم تمثيل بياني لمنحنى لورنز نقوم بنفس العمليات الحسابية السابقة والتي تبدأ بتحديد مراكز فئات المتغيرة (فئات الأجور)

الجدول رقم (6-02): توزيع عدد من العمال حسب أجورهم الشهرية (الوحدة : 10^3 دج)

	عدد العمال	فئات الأجور	مركز فئات الأجور	نسبة عمال كل فئة	نسبة مركز فئات الأجور	المتجمع الصاعد لنسب الأفراد	المتجمع الصاعد لنسب الأفراد
	20	70-50	60	20	12	20	12
	18	90-70	80	18	16	38	28
	20	110-90	100	20	20	58	48
	22	130-110	120	22	24	80	72
	20	150-130	140	20	28	100	100
المجموع	100	/	500	100	100	/	/

بعد إجراء هذه العمليات الحسابية نقوم بإجراء التمثيل البياني لمنحنى لورنز وفق الطريقة السابقة.

VI-3. مؤشر التمركز جيني:

VI-3-1. تعريف مؤشر التمركز: يعرف مؤشر التمركز بمعامل جيني (GINI Coefficient)، الذي ينسب إلى الاحصائي الإيطالي (كورادو جيني – Corrado Gini) [1] الذي يعتبر من أكثر الطرق المستخدمة لقياس التفاوت في توزيع الدخل أو متغيرة مدروسة أخرى، فهو يقيس مدى إبتعاد توزيع الدخل بين الأفراد في مجتمع ما عن خط المساواة ، أي أنه يقيس نسبة مساحة المنطقة بين منحنى لورنز وخط المساواة (المنطقة الملونة) من مساحة المثلث الموجود به هذه المنطقة، ففي حالة المساواة التامة ينطبق منحنى لورنز على خط المساواة فتصبح هذه النسبة معدومة (صفر) ، أما أقصى حالة من عدم المساواة في توزيع الدخل ، أي الحالة التي لا يحصل فيها أي من أفراد المجتمع على أي دخل بإستثناء فرد واحد يستحوذ على كل الدخل ففي هذه

[1] - عبد الرزاق عزوز ، الكامل في الإحصاء ، الجزء الثاني، ديوان المطبوعات الجامعية، الجزائر، 2011، ص 27.

الحالة ينطبق منحنى لورنز على من محور التجمع النسبي للأفراد والضلع الأيمن للمثلث تحت خط المساواة ، أي أن مساحة فجوة التوزيع تصبح مساحة المثلث ، وبهذا تصبح النسبة تساوي 1 أو 100 % .

VI-3-2. إستخراج مؤشر أو معامل التمركز (جيني): إن معامل التمركز (جيني) يتم إستخراجه من منحنى التمركز (لورنز) ، ولأجل توضيح ذلك ناخذ المثال الموالي :

مثال رقم (6–03) : تم توزيع مبلغ مالي قدره 145 وحدة نقدية على مجموعة أشخاص كما هو موضح في الجدول الموالي :

الجدول رقم (6–03): توزيع مبلغ مالي على أربعة أشخاص.

الشخص	محمد	أحمد	علي	سعيد	المجموع
المبلغ	25	30	40	50	145

المطلوب : ضع تمثيلا بيانيا يوضح توزيع المبلغ على الأشخاص إذا كان عادلا أم لا ؟

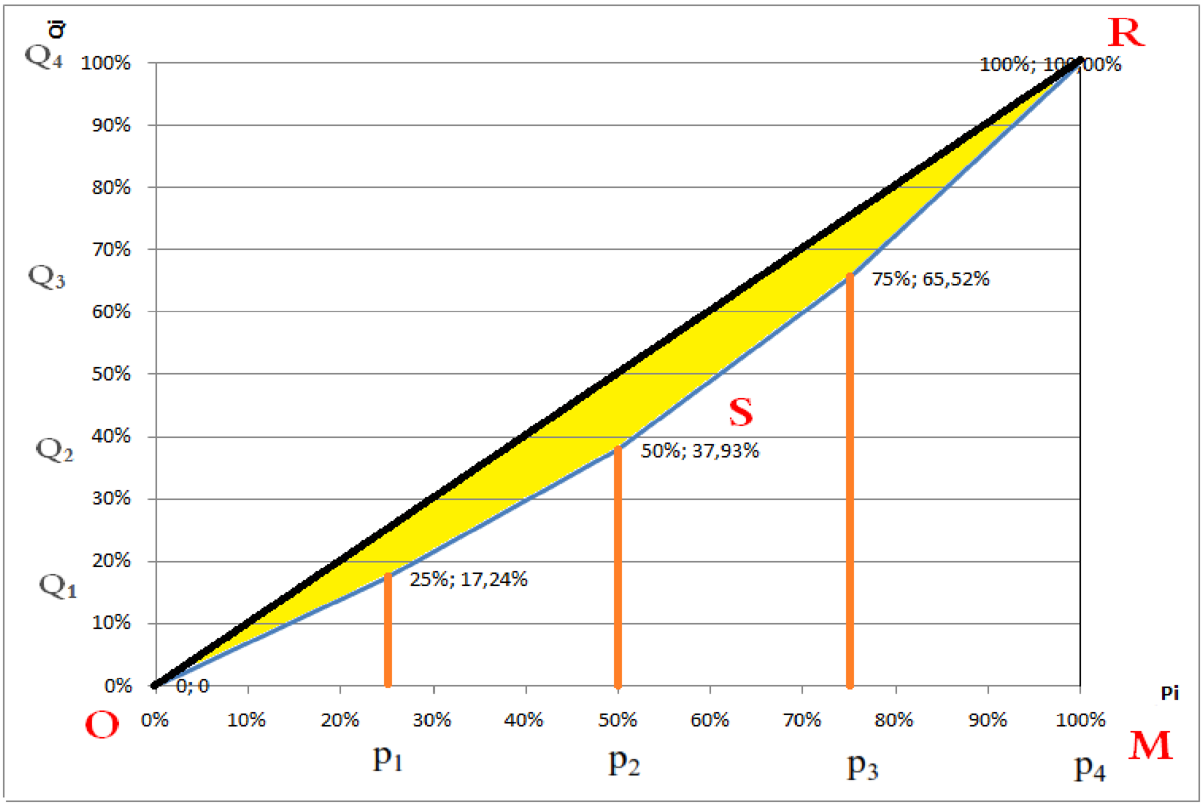
الحل: من أجل رسم التمثيل البياني لابد من حساب التوزيع التراكمي للأفراد والمبالغ

الجدول رقم (6–04): حساب التوزيعات التراكمية للأفراد والمبالغ.

الشخص X_i	المبلغ Y_i	نسبة الفرد	نسبة المبلغ	التوزيع التراكمي للأفراد P_i	التوزيع التراكمي للمبلغ Q_i
محمد	25	% 25	% 17.24	% 25	% 17.24
أحمد	30	% 25	% 20.69	% 50	% 37.93
علي	40	% 25	% 27.59	% 75	% 65.52
سعيد	50	% 25	% 34.48	% 100	% 100
المجموع	145	% 100	% 100	/	/

سنقوم بتمثيل بيانات الجدول في منحنى لورنز فكان ذلك موضح فيما يلي :

الشكل رقم (6-03) : التمثيل البياني الموضح لطريقة إستخراج معامل التمركز جيني إنطلاقا من منحنى التمركز (لورنز)



إن منحنى لورنز يبين الفجوة الموجودة في توزيع المبلغ (المنطقة الملونة) ويتم قياس هذه الفجوة أو المنطقة بإستخدام معامل جيني ، ويكون ذلك بقسمة المساحة الملونة (ORS) على مساحة المثلث (ORM) ، أي أن :

$$GINI = \frac{ORS}{ORM}$$

إن مساحة المثلث (ORM) يتم حسابها وفق قاعدة : $\frac{\text{الارتفاع x القاعدة}}{2} = \frac{100\% \ x \ 100\%}{2} = \frac{1}{2}$

أما المساحة الملونة (ORS) فهي مساحة المثلث (ORM) ناقص المساحة (OSRM) ، وبالتالي يصبح معامل جيني :

$$GINI = \frac{ORS}{ORM} = \frac{ORM - OSRM}{ORM} = 1 - \frac{OSRM}{ORM} = 1 - \frac{OSRM}{\frac{1}{2}} = 1 - 2(OSRM)$$

$$GINI = 1 - 2(OSRM)$$

والمساحة(OSRM) يتم حسابها كما يلي :

نلاحظ أن المساحة مقسمة على أربعة أقسام :

- القسم الأول هو مساحة مثلث قاعدته طولها P_1 وإرتفاع طوله Q_1 ، ومساحته يتم حسابها وفق قاعدة :

$$Z_1 = \frac{\text{الارتفاع x القاعدة}}{2} = \frac{P_1 \ * \ Q_1}{2}$$

- القسم الثاني فهو مستطيل طوله $(P_2 - P_1)$ وعرضه Q_1 وأعلاه مثلث قاعدته $(P_2 - P_1)$ وإرتفاع طوله $(Q_2 - Q_1)$ ، أي أن مساحة القسم الثاني هي :

$$Z_2 = \left(\text{العرض} * \text{الطول}\right) + \frac{\text{الارتفاع} * \text{القاعدة}}{2} = (P_2 - P_1) * Q_1 + \frac{(P_2 - P_1)\ x\ (Q_2 - Q_1)}{2}$$

$$Z_2 = \frac{2(P_2 - P_1) * Q_1}{2} + \frac{(P_2 - P_1)\ x\ (Q_2 - Q_1)}{2} = \frac{(P_2 - P_1)(2Q_1 + Q_2 - Q_1)}{2}$$

$$Z_2 = \frac{(P_2 - P_1)(Q_2 + Q_1)}{2}$$

- القسم الثالث فهو كذلك مستطيل طوله $(P_3 - P_2)$ وعرضه Q_2 وأعلاه مثلث قاعدته $(P_3 - P_2)$ وإرتفاع طوله $(Q_3 - Q_2)$ ، أي أن مساحة القسم الثالث هي :

$$Z_3 = \left(\text{العرض} * \text{الطول}\right) + \frac{\text{الارتفاع} * \text{القاعدة}}{2} = (P_3 - P_2) * Q_2 + \frac{(P_3 - P_2)\ x\ (Q_3 - Q_2)}{2}$$

$$Z_3 = \frac{2(P_3 - P_2) * Q_2}{2} + \frac{(P_3 - P_2)\ x\ (Q_3 - Q_2)}{2} = \frac{(P_3 - P_2)(2Q_2 + Q_3 - Q_2)}{2}$$

$$Z_3 = \frac{(P_3 - P_2)(Q_3 + Q_2)}{2}$$

- القسم الرابع فهو كذلك مستطيل طوله $(P_4 - P_3)$ وعرضه Q_3 وأعلاه مثلث قاعدته $(P_4 - P_3)$ وإرتفاع طوله $(Q_4 - Q_3)$ ، أي أن مساحة القسم الرابع هي :

$$Z_4 = \left(\text{العرض} * \text{الطول}\right) + \frac{\text{الارتفاع} * \text{القاعدة}}{2} = (P_4 - P_3) * Q_3 + \frac{(P_4 - P_3)\ x\ (Q_4 - Q_3)}{2}$$

$$Z_4 = \frac{2(P_4 - P_3) * Q_3}{2} + \frac{(P_4 - P_3)\ x\ (Q_4 - Q_3)}{2} = \frac{(P_4 - P_3)(2Q_3 + Q_4 - Q_3)}{2}$$

$$Z_4 = \frac{(P_4 - P_3)(Q_4 + Q_3)}{2}$$

نلاحظ أن مساحات الأقسام الأربعة تتشابه من حيث القاعدة ، حتى القسم الأول يمكن كتابة مساحته كما يلي :

$$Z_1 = \frac{P_1 * Q_1}{2} = \frac{(P_1 - P_0)(Q_1 + Q_0)}{2}$$

لأن $Q_0 = 0$ و $P_0 = 0$

إذن :

$$\text{GINI} = 1 - 2(\text{OSRM}) = 1 - 2(Z) = 1 - 2(Z_1 + Z_2 + Z_3 + Z_4)$$

$$\text{GINI} = 1 - 2(\text{OSRM}) = 1 - 2(Z) = 1 - 2\left(\sum_{i=1}^{4} \frac{(P_i - P_{i-1})(Q_i + Q_{i-1})}{2}\right)$$

$$\text{GINI} = 1 - \sum_{i=1}^{4} (P_i - P_{i-1})(Q_i + Q_{i-1})$$

إذن يمكننا الآن حساب معامل جيني لمثالنا السابق.

الجدول رقم (6-04): إستخدام جدول التوزيع في حساب معامل التمركز جيني.

الشخص X_i	المبلغ Y_i	نسبة الفرد	نسبة المبلغ	التوزيع التراكمي للأفراد P_i	التوزيع التراكمي للمبلغ Q_i	$(P_i - P_{i-1})$	$(Q_i + Q_{i-1})$	$(P_i - P_{i-1})(Q_i + Q_{i-1})$
محمد	25	% 25	% 17.24	% 25	% 17.24	25%	17,24%	0,0431
أحمد	30	% 25	% 20.69	% 50	% 37.93	25%	55,17%	0,137925
علي	40	% 25	% 27.59	% 75	% 65.52	25%	103,45%	0,258625
سعيد	50	% 25	% 34.48	% 100	% 100	25%	165,52%	0,4138
المجموع	145	% 100	% 100	/	/	/	/	0,85345

$$G = 1 - \sum_{i=1}^{4} (P_i - P_{i-1})(Q_i + Q_{i-1}) = 1 - 0.85345 = 0.14655$$

VI-3-3. خواص معامل التمركز جيني : يتميز معامل جيني بالخواص التالية :

- هو المقياس الأكثر شيوعا من حيث قياس عدالة توزيع الدخل القومي أو عدالة توزيع أي ظاهرة ؛
- يعطي قياسا رقميا لعدالة التوزيع دون إعطاء معلومات عن خصائص التوزيع مثل التموقع أو الشكل ؛
- هو نسبة تتراوح ما بين 0 و1 أو مابين 0 و 100% .

مثال رقم (6-04): قامت إحدى ولايات الجزائر بتوزيع أراضي على مجموعة من التعاونيات الفلاحية ، حيث أن كل تعاونية تحوز على مجموعة من الأشخاص ، وكان هذا التوزيع موضح في الجدول الموالي :

الجدول رقم (6-05): توزيع مساحات أراضي على مجموعة من أفراد تعاونيات.

رقم التعاونية	عدد أفراد التعاونية	مساحة الأرض (هكتار)
1	3	15
2	4	30
3	4	32
4	2	14
5	5	45
6	2	18
المجموع	20	154

المطلوب: هل ترى هناك عدالة في توزيع مساحة الأراضي من حيث عدد الأفراد ؟ برر ذلك بيانيا وحسابيا ؟

<u>الحل</u> :

1- حسب القراءة الأولية للأرقام نلاحظ أنه لا توجد عدالة في توزيع الأراضي لأنه مثلا عند تقسيم مساحة 15 هكتار على 3 أشخاص من التعاونية الأولى يأخذ كل شخص 5 هكتار، وأما في التعاونية الثانية يأخذ كل شخص 7.5 هكتار، ونفس الشيء في باقي التعاونيات، إلا أن التوزيع يقترب إلى التساوي.

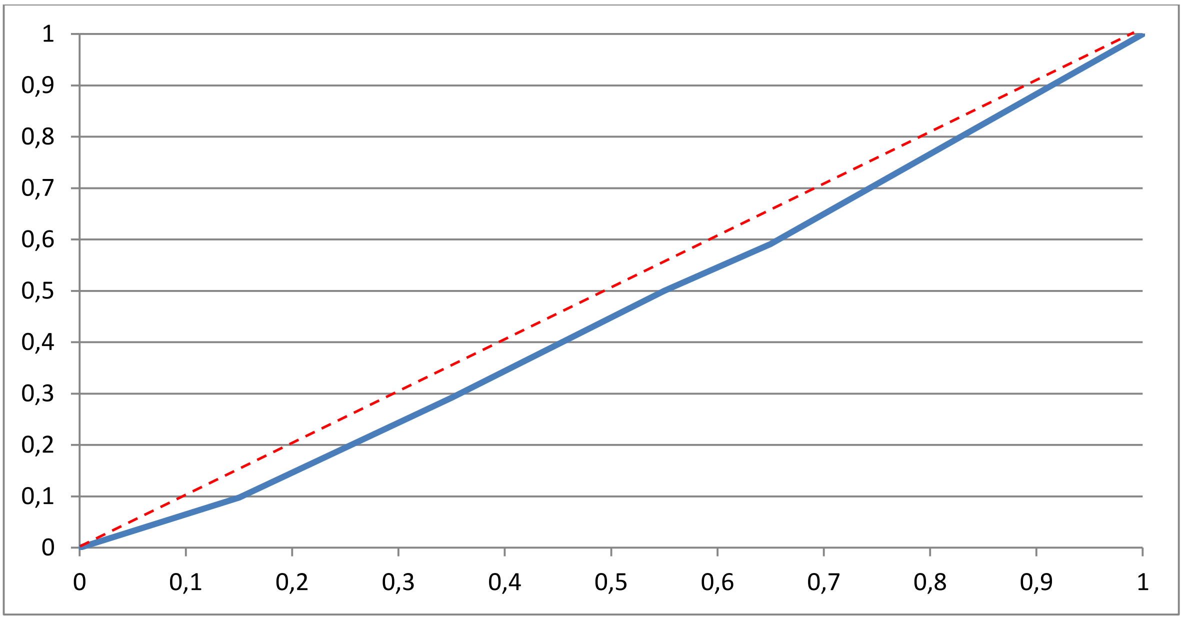
2- التأكد من وجود عدم مساواة في توزيع الأراضي من حيث عدد الأفراد وذلك بيانيا (منحنى لورنز) وحسابيا (معامل جيني):

أولا : التمثيل البياني لمنحنى لورنز والذي نستخدم فيه التوزيعات التراكمية للأفراد والأراضي.

الجدول رقم (6-06): المعطيات اللازمة لرسم منحنى لورنز وحساب معامل جيني

رقم التعاونية	عدد أفراد التعاونية X_i	مساحة الأرض (هكتار) Y_i	نسبة الأفراد	نسبة الأرض	التوزيع التراكمي للأفراد P_i	التوزيع التراكمي للأرض Q_i	$(P_i - P_{i-1})$	$(Q_i + Q_{i-1})$	$(P_i - P_{i-1})(Q_i + Q_{i-1})$
1	3	15	0,15	0,10	0,15	0,10	0,15	0,10	0,01
2	4	30	0,20	0,19	0,35	0,29	0,20	0,39	0,08
3	4	32	0,20	0,21	0,55	0,50	0,20	0,79	0,16
4	2	14	0,10	0,09	0,65	0,59	0,10	1,09	0,11
5	5	45	0,25	0,29	0,90	0,88	0,25	1,47	0,37
6	2	18	0,10	0,12	1	1	0,1	1,88	0,19
المجموع	20	154	1	1	/	/	/	/	0,92

الشكل رقم (6-04) : التمثيل البياني لمنحنى لورنز لمعطيات المثال رقم (6-04)



نلاحظ من منحنى التمركز لورنز أن خط توزيع الأراضي لا يبتعد كثيرا عن خط المساواة (الخط المتقطع) وبالتالي يمكن القول أن هناك نوع من المساواة في توزيع الأراضي على أفراد التعاونيات، ويمكن قياس ذلك بواسطة معامل التمركز جيني.

ثانيا : حساب معامل جيني وفق المعطيات المحسوبة في الجدول السابق

$$G = 1 - \sum_{i} (P_i - P_{i-1})(Q_i + Q_{i-1}) = 1 - 0.92 = 0.08$$

المحور السابع:

الأرقام القياسية

VII-1. تمهيد:

تُعدّ الأرقام القياسية من أهم الأدوات الإحصائية التي تُستخدم لقياس التغيرات التي تطرأ على ظواهر اقتصادية واجتماعية عبر الزمن. فالأسعار والكميات ومستويات الإنتاج والاستهلاك كلها متغيرات تتغير باستمرار، ويحتاج الباحث وصانع القرار إلى وسيلة دقيقة وبسيطة لفهم اتجاهاتها وتقييم حجم التغير فيها. وهنا يأتي دور الأرقام القياسية التي تختزل معلومات معقّدة في قيمة عددية واحدة تعبّر عن مستوى التغير بين فترتين أو أكثر.

كما تكتسب الأرقام القياسية أهميتها لكونها أساسًا تعتمد عليه الحكومات والمؤسسات في تحليل الظواهر الاقتصادية مثل التضخم، تكاليف المعيشة، قوة الإنتاج، والتغير في الطلب والعرض. وهي توفر مؤشرات تساعد على مقارنة الفترات الزمنية، ودراسة الاتجاهات طويلة المدى مما يجعلها عنصرًا أساسيًا في التخطيط ووضع السياسات.

VII-2. مفهوم الرقم القياسي:

يعرف الرقم القياسي على أنه مؤشر أو قيمة تعبر عن المستوى العام للتغير في قيم ظاهرة معينة خلال فترات أو أماكن معينة، ويستخدم لمقارنة التغير في ظاهرة واحدة، ولمقارنة التغير في المستوى العام لمجموعة من المتغيرات أو الظواهر المختلفة فيما بينها والمرتبطة بشكل أو بآخر لتكون مجموعة متجانسة، إذن الرقم القياسي هو مقياس إحصائي يبين التغيرات التي تطرأ على ظاهرة ما، أو على مجموعة من الظواهر المترابطة خلال فترة زمنية أو منطقة جغرافية [1].

VII-3. أنواع الأرقام القياسية:

للأرقام القياسية عدة أنواع، من أهمها:

VII-3-1. الرقم القياسي البسيط:

يسمى كذلك بالرقم القياسي الفردي أو البدائي، وهو رقم قياس لسلعة واحدة فقط، ويكون ذلك إما لسعرها أو كميتها أو قيمتها.

VII-3-1-1. الرقم القياسي البسيط للسعر:

يرمز له بالرمز (I_P) ، ويحسب بالصيغة التالية:

$$I_{P(t_1/t_0)} = \frac{P_1}{P_0} \times 100$$

حيث : P_0 : سعر السلعة في سنة الأساس t_0

P_1 : سعر السلعة في سنة المقارنة t_1

VII-3-1-2. الرقم القياسي البسيط للكمية:

يرمز له بالرمز (I_Q) ، ويحسب بالصيغة التالية:

[1] - عدنان عباس حميدان وآخرون، مرجع سابق، ص ص : 362، 363

$$I_{Q\left(t_1/t_0\right)} = \frac{Q_1}{Q_0} \times 100$$

حيث: Q_0 : كمية السلعة في سنة الأساس t_0

Q_1 : كمية السلعة في سنة المقارنة t_1

3-1-3-VII. الرقم القياسي البسيط للقيمة:

هو رقم قياسي للجودة، أي أن كل سلعة ترتفع قيمتها يكون بسبب جودتها، يرمز له بالرمز (I_V) ، ويحسب بالصيغة التالية:

$$I_{V\left(t_1/t_0\right)} = \frac{V_1}{V_0} \times 100 = \frac{P_1 Q_1}{P_0 Q_0} \times 100$$

حيث: V_0 : قيمة السلعة في سنة الأساس t_0

V_1 : قيمة السلعة في سنة المقارنة t_1

مثال رقم (7-01) : عرفت البطاطا في سوق ما خلال الفترة (2018-2023) تغيرات هامة سواءا من حيث السعر أو الكمية المباعة كما هو مبين في الجدول الموالي:

الجدول رقم (7-01): تغيرات سعر البطاطا والكمية المباعة منها خلال الفترة (2018-2023)

السنة	2018	2019	2020	2021	2022	2023
السعر (دج)	40	45	45	50	55	60
الكمية المباعة (كغ)	60	62	66	68	70	72

المطلوب: حساب تغيرات سعر البطاطا وكميتها المباعة، وقيمتها باعتبار أن سنة 2018 هي سنة الأساس؟

الحل:

نقوم بحساب تغيرات السعر والكمية المباعة والقيمة اعتمادا على الجدول الموالي:

الجدول رقم (7-02): حسابات الأرقام القياسية البسيطة باستخدام العلاقات السابقة

المعطيات				حسابات الأرقام القياسية		
السنة	السعر (دج)	الكمية المباعة (كغ)	القيمة	للسعر	للكمية	للقيمة
2018	40	60	2400	100	100	100
2019	45	63	2835	112,5	105	118,125
2020	45	66	2970	112,5	110	123,75
2021	50	69	3450	125	115	143,75
2022	55	72	3960	137,5	120	165
2023	60	75	4500	150	125	187,5

من الأرقام القياسية المحسوبة في الجدول نستنتج ما يلي:

- ارتفع سعر البطاطا بنسبة 50% خلال الفترة 2018-2023 وتحسب النسبة من الجدول من خلال :

$$150-100=50\%$$

- ارتفعت الكمية المباعة من البطاطا بنسبة 25% خلال الفترة 2018-2023 وتحسب النسبة من الجدول من خلال :

$$125-100=25\%$$

- ارتفعت قيمة البطاطا بنسبة 87.5% خلال الفترة 2018-2023 وتحسب النسبة من الجدول من خلال :

$$187{,}5-100=87{,}5\%$$

VII-3-2. الرقم القياسي المركب:

ينقسم إلى قسمين، الرقم القياسي التجميعي والرقم القياسي المرجح.

VII-3-2-1. الرقم القياسي التجميعي:

يستخدم هذا المقياس لمعرفة التغيرات التي حدثت لمجموعة من السلع بأسعارها الحقيقية خلال فترتين، سواءا من حيث السعر أو الكمية أو القيمة، ويحسب الرقم القياسي التجميعي حسب الصيغ التالية:

VII-3-2-1-1. الرقم القياسي التجميعي للأسعار:

يرمز له بالرمز (IC_P) ، ويحسب بالصيغة التالية:

$$IC_{P(t_1/t_0)}=\frac{\sum P_1}{\sum P_0}\times 100$$

حيث: $\sum P_0$ هو مجموع أسعار السلع في سنة الأساس t_0

$\sum P_1$ هو مجموع أسعار السلع في سنة المقارنة t_1

VII-3-2-1-2. الرقم القياسي التجميعي للكميات:

يرمز له بالرمز (IC_Q) ، ويحسب بالصيغة التالية:

$$IC_{Q(t_1/t_0)}=\frac{\sum Q_1}{\sum Q_0}\times 100$$

حيث: $\sum Q_0$ هو مجموع كميات السلع في سنة الأساس t_0

$\sum Q_1$ هو مجموع كميات السلع في سنة المقارنة t_1

VII-3-2-1-3. الرقم القياسي التجميعي للقيم:

يرمز له بالرمز (IC_V) ، ويحسب بالصيغة التالية:

$$IC_{V\left(t_1/t_0\right)} = \frac{\sum V_1}{\sum V_0} \times 100 = \frac{\sum P_1 Q_1}{\sum P_0 Q_0} \times 100$$

حيث: $\sum V_0$ هو مجموع قيم السلع في سنة الأساس t_0

$\sum V_1$ هو مجموع قيم السلع في سنة المقارنة t_1

مثال رقم (7–02) : عرفت البطاطا والبصل والجزر تغيرات هامة سواءا من حيث السعر أو الكمية المباعة خلال عام 2023 بالمقارنة بعام 2018 كما هو موضح في الجدول الموالي:

الجدول رقم (7–03): تغيرات سعر مجموعة من الخضر والكمية المباعة منها خلال سنتي 2018 و 2023

الخضر	2018		2023	
	السعر (دج)	الكمية المباعة (كغ)	السعر (دج)	الكمية المباعة (كغ)
البطاطا	40	60	60	75
البصل	20	40	40	65
الجزر	50	30	80	40

المطلوب: حساب تغيرات سعر الخضر وكميتها المباعة، وقيمتها باعتبار أن سنة 2018 هي سنة الأساس؟

الحل: سنقوم بحساب قيم الخضر

الجدول رقم (7–04): حساب قيم الخضر ومجموع الأسعار ومجموع الكميات ومجموع القيم.

الخضر	السعر (دج)		الكمية المباعة (كغ)		القيم	
	2018	2023	2018	2023	2018	2023
البطاطا	40	60	60	75	2400	4500
البصل	20	40	40	65	800	2600
الجزر	50	80	30	40	1500	3200
المجموع $\sum$	110	180	130	180	4700	10300

– الرقم القياسي التجميعي للأسعار:

$$IC_{P\left(2023/2018\right)} = \frac{\sum P_{2023}}{\sum P_{2018}} \times 100 = \frac{180}{110} \times 100 = 163{,}63\%$$

إن أسعار الخضر ارتفعت بنسبة 63.63% سنة 2023 بالمقارنة بسنة الأساس 2018.

– الرقم القياسي التجميعي للكميات المباعة:

$$IC_{Q(2023/2018)} = \frac{\sum Q_{2023}}{\sum Q_{2018}} \times 100 = \frac{180}{130} \times 100 = 138{,}46\%$$

إن الكمية المباعة من الخضر ارتفعت بنسبة 38.46% سنة 2023 بالمقارنة بسنة الأساس 2018.

- الرقم القياسي التجميعي للقيم:

$$IC_{V(2023/2018)} = \frac{\sum V_{2023}}{\sum V_{2018}} \times 100 = \frac{\sum P_{2023}Q_{2023}}{\sum P_{2018}Q_{2018}} \times 100 = \frac{10300}{4700} \times 100 = 219{,}14\%$$

إن قيمة الخضر ارتفعت بنسبة 119.14% سنة 2023 بالمقارنة بسنة الأساس 2018.

VII-3-2-2. الرقم القياسي المرجح:

إن من بين عيوب الرقم القياسي التجميعي هو أنه لا يحدد وزن كل سلعة وأهميتها النسبية داخل المجموعة، وبالتالي جاءت الصيغ المرجحة للأرقام القياسية للتغلب على هذه العيوب، ومن أبرز الأرقام القياسية المرجحة وأكثرها استخداما:

VII-3-2-2-1. الرقم القياسي المرجح للاسبير:

يرى لاسبير أن معطيات سنة الأساس هي الأهم، وبالتالي يتم ترجيحها سواءا للأسعار أو الكميات كما هو موضح في الصيغتين التاليتين:

- الرقم القياسي المرجح للأسعار للاسبير (IL_P) يكون كما يلي:

$$IL_{P(t_1/t_0)} = \frac{\sum P_1Q_0}{\sum P_0Q_0} \times 100$$

- الرقم القياسي المرجح للكميات للاسبير (IL_Q) يكون كما يلي:

$$IL_{Q(t_1/t_0)} = \frac{\sum Q_1P_0}{\sum Q_0P_0} \times 100$$

مثال رقم (7-03) : نأخذ معطيات المثال رقم (7-02) السابق

الجدول رقم (7-05): تغيرات سعر مجموعة من الخضر والكمية المباعة منها خلال سنتي 2018 و 2023

الخضر	2018		2023	
	السعر (دج)	الكمية المباعة (كغ)	السعر (دج)	الكمية المباعة (كغ)
البطاطا	40	60	60	75
البصل	20	40	40	65
الجزر	50	30	80	40

المطلوب: باستخدام الرقم القياسي للاسبير أحسب تغيرات سعر الخضر وكميتها المباعة باعتبار أن سنة 2018 هي سنة الأساس؟

الحل:

سنستخدم الجدول السابق في حساب الأرقام القياسية للاسبير

الجدول رقم (7-06): حساب معطيات الأرقام القياسية للاسبير

الخضر	السعر (دج)		الكمية المباعة (كغ)		القيم		
	2018	2023	2018	2023	$P_{2023} \times Q_{2018}$	$P_{2018} \times Q_{2018}$	$Q_{2023} \times P_{2018}$
البطاطا	40	60	60	75	3600	2400	3000
البصل	20	40	40	65	1600	800	1300
الجزر	50	80	30	40	2400	1500	2000
المجموع	/	/	/	/	7600	4700	6300

- الرقم القياسي المرجح للاسبير للأسعار:

$$IL_{P(2023/2018)} = \frac{\sum P_{2023}Q_{2018}}{\sum P_{2018}Q_{2018}} \times 100 = \frac{7600}{4700} \times 100 = 161{,}7\%$$

- الرقم القياسي المرجح للاسبير للكميات:

$$IL_{Q(2023/2018)} = \frac{\sum Q_{2023}P_{2018}}{\sum Q_{2018}P_{2018}} \times 100 = \frac{6300}{4700} \times 100 = 134{,}04\%$$

نلاحظ أن الأسعار إرتفعت بنسبة 61.7% سنة 2023 بالمقارنة بسنة الأساس 2018، أما الكميات المباعة فقد ارتفعت بنسبة 34.04% عام 2023 بالمقارنة بسنة الأساس 2018، وهي نسب أقل من المحسوبة بواسطة الرقم القياسي التجميعي.

VII-3-2-2-2. الرقم القياسي المرجح لباش:

يرى باش أن معطيات سنة المقارنة هي الأهم، وهذا عكس لاسبير، وبالتالي يتم ترجيح سنة المقارنة سواءا للأسعار أو الكميات كما هو موضح في الصيغتين التاليتين:

- الرقم القياسي المرجح للأسعار لباش (IP_P) يكون كما يلي:

$$IP_{P(t_1/t_0)} = \frac{\sum P_1Q_1}{\sum P_0Q_1} \times 100$$

- الرقم القياسي المرجح للأسعار لباش (IP_Q) يكون كما يلي:

$$IP_{Q(t_1/t_0)} = \frac{\sum Q_1P_1}{\sum Q_0P_1} \times 100$$

مثال رقم (7-04) : نأخذ معطيات المثال رقم (7-02) السابق

المطلوب: باستخدام الرقم القياسي لباش أحسب تغيرات سعر الخضر وكميتها المباعة باعتبار أن سنة 2018 هي سنة الأساس؟

الجدول رقم (7–07): حساب معطيات الأرقام القياسية لباش

الخضر	السعر (دج)		الكمية المباعة (كغ)		القيم		
	2018	2023	2018	2023	$P_{2023}\times Q_{2023}$	$P_{2018}\times Q_{2023}$	$Q_{2018}\times P_{2023}$
البطاطا	40	60	60	75	4500	3000	3600
البصل	20	40	40	65	2600	1300	1600
الجزر	50	80	30	40	3200	2000	2400
المجموع	/	/	/	/	10300	6300	7600

– الرقم القياسي المرجح لباش للأسعار:

$$IP_{P(2023/2018)} = \frac{\sum P_{2023}Q_{2023}}{\sum P_{2018}Q_{2023}}\times 100 = \frac{10300}{6300}\times 100 = 163{,}49\%$$

– الرقم القياسي المرجح لباش للكميات:

$$IP_{Q(2023/2018)} = \frac{\sum Q_{2023}P_{2023}}{\sum Q_{2018}P_{2023}}\times 100 = \frac{10300}{7600}\times 100 = 135{,}52\%$$

نلاحظ أن الأسعار ارتفعت بنسبة 63.49% سنة 2023 بالمقارنة بسنة الأساس 2018، أما الكميات المباعة فقد ارتفعت بنسبة 35.52% عام 2023 بالمقارنة بسنة الأساس 2018، وهي نسب أكبر من المحسوبة بواسطة الرقم القياسي المرجح للاسبير.

VII-3-2-2-3. الرقم القياسي المرجح لفيشر:

إعتمد فيشر على الرقمين القياسيين للاسبير وباش معا من خلال متوسط هندسي لهما وفق الصيغة التالية:

$$IF_{(t_1/t_0)} = \sqrt{IL \times IP}\times 100$$

أما الأرقام القياسية للأسعار والكميات فتكون كما يلي:

- الرقم القياسي المرجح للأسعار لفيشر (IF_P) يكون كما يلي:

$$IF_{P(t_1/t_0)} = \sqrt{\frac{\sum P_1Q_0}{\sum P_0Q_0}\times\frac{\sum P_1Q_1}{\sum P_0Q_1}}\times 100$$

- الرقم القياسي المرجح للكميات لفيشر (IF_Q) يكون كما يلي:

$$IF_{Q(t_1/t_0)} = \sqrt{\frac{\sum Q_1P_0}{\sum Q_0P_0}\times\frac{\sum Q_1P_1}{\sum Q_0P_1}}\times 100$$

مثال رقم (7–05) : نواصل مع معطيات المثال رقم (7–02) السابق

المطلوب: باستخدام الرقم القياسي لفيشر أحسب تغيرات سعر الخضر وكميتها المباعة باعتبار أن سنة 2018 هي سنة الأساس؟

الحل:

من خلال ما توصلنا إليه سابقا وهو أن:

$$IL_{P(2023/2018)} = 161,7\%$$

$$IL_{Q(2023/2018)} = 134,04\%$$

$$IP_{P(2023/2018)} = 163,49\%$$

$$IP_{Q(2023/2018)} = 135,52\%$$

إذن:

- الرقم القياسي المرجح للأسعار لفيشر هو:

$$IF_{P(t_1/t_0)} = \sqrt{IL_P \times IP_P} \times 100 = \sqrt{1,617 \times 1,6349} \times 100 = \sqrt{2,6436} \times 100 = 162,59\%$$

- الرقم القياسي المرجح للكميات لفيشر هو:

$$IF_{Q(t_1/t_0)} = \sqrt{IL_Q \times IP_Q} \times 100 = \sqrt{1,3404 \times 1,3552} \times 100 = \sqrt{1,8165} \times 100 = 134,77\%$$

المحور الثامن:

الارتباط والانحدار

VIII-1. تمهيد:

لقد عرضنا من خلال المحاور السابقة طرق التعامل مع متغيرة واحدة بداية من تحديد نوعها (كمية أو نوعية) إلى غاية تحديد شكل توزيعها (مدبب أو مفرطح) ، لكن في الواقع قد نتعامل مع أكثر من متغيرة ، فمثلا إذا ذهبت للطبيب من أجل العلاج من السمنة فلا يقوم بتحديد وزنك فقط بل يربطه بقامتك ، فلو كان طولك 1.75 م يطلب منك إنقاص وزنك إلى المجال [70 75] كلغ ، كما أن الفرد عادة ما يقوم بتقييد إستهلاكه (مصاريفه) إنطلاقا من دخله المتاح، وبالتالي دراسة متغيرة دون أخرى يعتبر إهمالا للعلاقة أو التأثير المتبادل بين المتغيرتين أو حتى تغير بيانات المتغيرة في حد ذاتها.

VIII-2. المتغيرات الثنائية (ثنائية التغير)

إن المتغيرات الثنائية هي متغيرتين بتكرار مشترك ، وهذا ما صادفناه في الجداول التكرارية المزدوجة ، فهذه الجداول تتكون من بيانات متغيرة في السطر وبيانات متغيرة أخرى في العمود وبينهما تكرارات مشتركة ، كما يسمى هذا النوع من الجداول بجداول التوافق ، فمن مميزات هاتين المتغيرتين أنهما تتغيران بقيم مشتركة.

مثال رقم (8-01): من أجل دراسة العلاقة ما بين دخل الأسرة وعدد أفرادها تم الحصول على بيانات 20 عائلة، حيث:

- العائلات التي دخلها يتراوح ما بين 10 و 19 ألف دينار : يوجد 4 عائلات لها طفل و 3 عائلات لها طفلان ، 2 عائلات لها 3 أطفال ؛
- العائلات التي دخلها يتراوح ما بين 20 و 29 ألف دينار : يوجد 5 عائلات لها طفل و 6 عائلات لها طفلان ، 7 عائلات لها 3 أطفال ؛
- العائلات التي دخلها يتراوح ما بين 30 و 39 ألف دينار : يوجد 6 عائلات لها طفل و 7 عائلات لها طفلان ، 8 عائلات لها 3 أطفال ؛

إن البيانات السابقة يمكن تبويبها في جدول تكراري كما يلي:

الجدول رقم (8-01): جدول توزيع تكراري مزدوج للمتغيرتين عدد أطفال العائلة ودخلها.

عدد أطفال العائلة / دخل العائلة	1	2	3	المجموع
من 10 إلى 19 ألف دينار	4	3	2	9
من 20 إلى 29 ألف دينار	5	6	7	18
من 30 إلى 39 ألف دينار	6	7	8	21
المجموع	15	16	17	48

من خلال قراءة أرقام الجدول يظهر أن هناك علاقة إرتباط ما بين دخل العائلة وعدد أطفالها ، فكلما زاد دخل العائلة زاد عدد أطفالها ، فمثلا بعدما صادفنا 9 عائلات يقل دخلها عن 19 ألف دينار لها ما بين طفل إلى 3 أطفال، إرتفع العدد إلى 18 عائلة التي دخلها ما بين 20 إلى 29 ألف دينار ، وبالمقابل عندما أصبح الدخل يتراوح ما بين 30 إلى 39 ألف دينار إرتفع العدد إلى 21 عائلة ، وبقراءة أخرى عدد الأسر التي دخلها يقل عن 19 ألف دينار ولها 3 أطفال يوجد عائلتين فقط ، أما بالمقابل فالعائلات التي دخلها يتراوح ما بين 30 إلى 39 ألف دينار ولها 3 أطفال عددها 8 عائلات.

كما قد نصادف ظاهرتين لا تحتويان على تكرارات مشتركة لكن لهما علاقة ترابط مثل علاقة الاستهلاك بالدخل المتاح والتي تسمى بدالة الاستهلاك $C = f(y)$ ، أي كلما تغير الدخل يتغير الاستهلاك ، فيسمى متغير الدخل بالمتغير المؤثر أو المتغير المستقل وعادة ما يرمز له بالرمز X_i ، أما متغير الاستهلاك فهو متغير متأثر أو المتغير التابع وعادة ما يرمز له بالرمز Y_i .

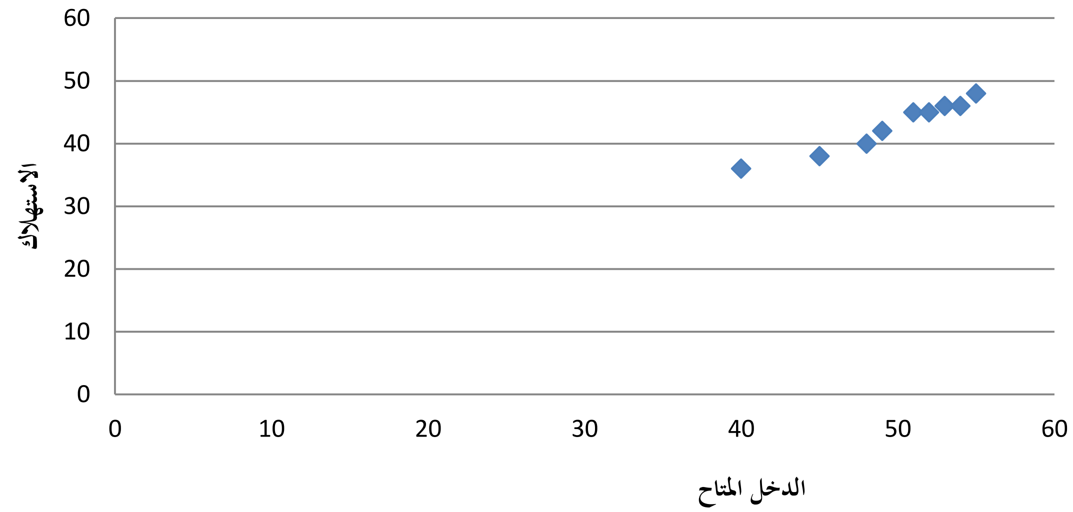
مثال رقم (8–02) : قدمت إحدى العائلات بيانات عن إستهلاكها (مصاريفها) ودخلها المتاح خلال الفترة 2010-2022، فكانت البيانات مدونة في الجدول الموالي :

الجدول رقم (8–02): بيانات عن إستهلاك إحدى العائلات ودخلها المتاح الوحدة : 10^3 دينار

السنة	2010	2011	2012	2013	2014	2015	2016	2017	2018
الدخل المتاح X_i	40	45	48	49	51	53	55	52	54
الاستهلاكY_i	36	38	40	42	45	46	48	45	46

من أجل توضيح علاقة المتغيرين يمكننا تمثيل أرقام الجدول بيانيا كالتالي :

الشكل رقم (8–01): شكل العلاقة بين الدخل المتاح والاستهلاك



نلاحظ أنه كلما زاد الدخل زاد الاستهلاك عند هذه الأسرة ، وكلما إنخفض الدخل إنخفض الاستهلاك ، أي أن إستهلاك هذه الأسرة يتأثر بدخلها المتاح ، أي أن هناك علاقة إرتباط بين المتغيرين.

VIII–3. الارتباط بين المتغيرات

إن معرفة وقياس الارتباط بين المتغيرات الاقتصادية يختلف حسب نوع هذه المتغيرات، فقد تكون هذه المتغيرات كمية أو نوعية.

VIII–3–1. الارتباط بين متغيرين كيفيين : هناك عدة معاملات لقياس الارتباط بين متغيرين كيفيين أبرزها معامل الاقتران فآي (Phi) ، كاي مربع (χ^2) ومعامل التوافق.

VIII–3–1–1. معامل الاقتران فآي (Phi) : يستخدم هذا المعامل في البحث عن إيجاد العلاقة بين متغيرين كيفيين كل منهما ثنائي الانقسام مثل وضعية الطالب (X) (حاضر أو غائب) ، ووضعه الصحي (Y) (مريض أو غير مريض) ، حيث يرمز لمعامل الاقتران (Phi) بالرمز $r_{\emptyset}$ ، ويتم حسابه كما يلي :

إذا كانت البيانات المبينة في الجدول الموالي والتي تمثل وضعية مجموعة من الطلبة من خلال الغياب أو الحضور ووضعهم الصحي :

الجدول رقم (8-03): جدول توزيع تكراري مزدوج لمتغيرتي وضعية الطالب من خلال الحضور أو الغياب ووضعه الصحي

		وضعية الطالب	
		غائب	حاضر
الوضع الصحي	مريض	a	b
للطالب	غير مريض	c	d

معامل الاقتران فآي (Phi) يحسب رياضيا كما يلي :

$$r_{\emptyset} = \frac{ad - bc}{ad + bc}$$

حيث : a ، b ، c ، d هي التكرارات المشتركة ، فمثلا a تمثل عدد من الطلبة الغائبين وفي نفس الوقت هم مرضى .

خواص معامل الاقتران فآي (Phi) : يتميز بما يلي

- يقيس معامل الاقتران فآي (Phi) العلاقة بين المتغيرات الكيفية ثنائية الانقسام ؛
- قيمة معامل الاقتران فآي (Phi) محصور في المجال [1- 1] ؛
- إذا كانت قيمة معامل الاقتران فآي (Phi) سالبة يدل ذلك على وجود علاقة عكسية (سالبة) بين المتغيرين ؛
- إذا كانت قيمة معامل الاقتران فآي (Phi) موجبة يدل ذلك على وجود علاقة طردية (موجبة) بين المتغيرين ؛
- إذا إقترب معامل الاقتران فآي (Phi) إلى (1-) أو (1) فإن ذلك يدل على أن العلاقة قوية بين المتغيرين ؛
- إذا إقترب معامل الاقتران فآي (Phi) إلى (0) فإن ذلك يدل على أن العلاقة ضعيفة بين المتغيرين ؛
- إذا كان a=0 أو d=0 فإن $r_{\emptyset} = -1$ ، أي توجد علاقة عكسية تامة ما بين المتغيرين ؛
- إذا كان c=0 أو b=0 فإن $r_{\emptyset} = 1$ ، أي توجد علاقة موجبة تامة ما بين المتغيرين ؛
- إذا كان a=0 أو d=0 و c=0 أو b=0 فإن $r_{\emptyset} = 0$ ، أي عدم وجود علاقة إقتران بين المتغيرين.

مثال رقم (8-03): البيانات الموالية تخص 35 شخص يراد دراسة العلاقة ما بين مستواهم التعليمي وطموحهم في الحياة.

الجدول رقم (8-04): توزيع 35 شخص حسب طموحهم في الحياة ومستواهم التعليمي.

المجموع	منخفض	مرتفع	الطموح / المستوى التعليمي
12	7	5	يوجد مستوى
23	8	15	لا يوجد مستوى
35	15	20	المجموع

المطلوب : ما هي قوة وإتجاه العلاقة بين المتغيرين (المستوى التعليمي و الطموح في الحياة) ؟

الحل:

$$r_{\emptyset} = \frac{ad - bc}{ad + bc} = \frac{5\text{x}8 - 7\text{x}15}{5\text{x}8 + 7\text{x}15} = \frac{-65}{145} = -0.45$$

وجدنا معامل الاقتران سالب و هو يدل على أن العلاقة بين المتغيرين عكسية ، وفي نفس الوقت النسبة أقل من 0.5 وبالتالي يمكن القول أن العلاقة بين المتغيرين ليست قوية.

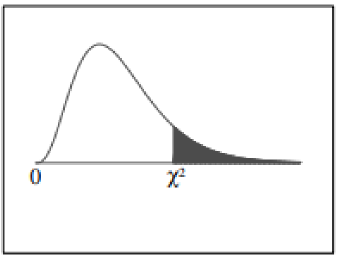
VIII–3–1–2. قيمة كاي مربع بدلالة معامل الاقتران $r_{\emptyset}$: في بعض الحالات لما تقترب قيمة معامل الاقتران من الصفر نقع في حالة شك من وجود علاقة أصلا أم لا ما بين المتغيرين ، وبالتالي نقوم بتأكيد ذلك من خلال حساب قيمة كاي مربع (χ_c^2) وفق القاعدة التالية :

$$\chi_c^2 = \emptyset^2 N$$

ثم يتم مقارنتها بقيمة مجدولة (χ_t^2) عند مستوى الدلالة المعنوية 5 % ودرجة الحرية (1) ، المبينة في الجدول الموالي :

الجدول رقم (8–05): جدول توزيع كاي مربع.

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

والقرار يكون كالتالي:

- إذا كانت القيمة المحسوبة أكبر من المجدولة $(\chi_c^2 > \chi_t^2)$ نقول أنه يوجد إرتباط ذو دلالة معنوية بين المتغيرين ؛
- إذا كانت القيمة المحسوبة أقل من المجدولة $\chi_c^2 < \chi_t^2$ نقول أنه لا يوجد إرتباط ذو دلالة معنوية بين المتغيرين ؛

مثال رقم (8–04) : تم الحصول على بيانات خاصة بطلبة إحدى الجامعات حول ركوبهم حافلة نقل الطلبة من عدمه فكانت النتائج كالتالي :

الجدول رقم (8–06): جدول توزيع تكراري مزدوج لمتغيرتي الركوب في الحافلة ونوع جنس الراكب.

الركوب في الحافلة / الجنس	يركب	لا يركب	المجموع
ذكر	4	2	6
أنثى	3	2	5
المجموع	7	4	11

المطلوب : تحديد قوة العلاقة بين جنس الطالب و ركوبه حافلة نقل الطلبة أو عدم ركوبه وتأكيد تلك العلاقة ؟

الحل :

$$r_{\emptyset} = \frac{ad - bc}{ad + bc} = \frac{4x2 - 2x3}{4x2 + 2x3} = \frac{2}{14} = 0.14$$

نلاحظ أن قيمة $r_{\emptyset}$ تقترب من الصفر وهذا يدل على أن العلاقة بين المتغيرين قليلة جدا ، نتأكد من دلالتها الاحصائية بواسطة إختبار كاي مربع ، حيث :

$$\chi_c^2 = \emptyset^2 N = (0.14)^2(11) = 0.2156$$

ولدينا :

$$\chi_{t(0.05\,;\,1)}^2 = 3.841$$

نلاحظ أن $\chi_t^2 > \chi_c^2$ ومنه يمكن نقول أنه لا يوجد إرتباط ذو دلالة معنوية بين المتغيرين (جنس الطالب وركوبه من عدمه في حافلة النقل)

VIII–3–1–3. معامل التوافق (*Coefficient of Contingency*): يستخدم هذا المعامل من أجل دراسة العلاقة بين متغيرين كيفيين لهما أكثر من إنقسامين (مثل تقديرات النجاح : مقبول ، حسن ، جيد ...) ، ويحسب هذا المعامل بالعلاقة الرياضية التالية :

$$C = \sqrt{\frac{F - 1}{F}}$$

حيث : F هو مجموع مربع كل خانة (تكرار مشترك) قسمة (مجموع عمودها ضرب مجموع صفها)

خواص معامل التوافق : يتميز معامل التوافق بما يلي

- يقيس معامل التوافق C قوة العلاقة بين المتغيرين النوعيين وليس إتجاه العلاقة ؛
- الحد الأدنى لـ C هو (0) ؛
- الحد الأعلى لـ C يتم حسابه كما يلي :

$$C_{Max} = \sqrt{\frac{L - 1}{2}}$$

حيث : L يمثل العدد الأقل من بين عدد الصفوف أو الأعمدة للمتغيرين (أي العدد الأقل من تقسيمات المتغيرين).

والقرار يكون كما يلي :

- إذا كانت C قريبة من C_{Max} نقول أن العلاقة بين المتغيرين قوية ؛
- إذا كانت C بعيدة من C_{Max} نقول أن العلاقة بين المتغيرين ضعيفة ؛

مثال رقم (8–05) : قام أحد الباحثين بجمع بيانات من عينة أشخاص في مدينة ما من حيث الشهادات التي تحصلوا عليها، فكانت النتائج كالتالي:

الجدول رقم (8–07): جدول توزيع تكراري مزدوج لمتغيرتي شهادات عينة أشخاص وجنسهم

الشهادة / الجنس	بكالوريا	ليسانس	ماستر	المجموع
ذكر	**3**	**5**	**12**	**20**
أنثى	**2**	**8**	**15**	**25**
المجموع	**5**	**13**	**27**	**45**

المطلوب : قم بقياس العلاقة بين جنس الشخص والشهادة التي تحصل عليها ؟

الحل : نقوم بدراسة قوة العلاقة بإستخدام معامل التوافيق

$$C=\sqrt{\frac{F-1}{F}}$$

$$F=\frac{3^2}{5x20}+\frac{5^2}{13x20}+\frac{12^2}{27x20}+\frac{2^2}{5x25}+\frac{8^2}{13x25}+\frac{15^2}{27x25}=1.015$$

$$C=\sqrt{\frac{1.015-1}{1.015}}=0.121$$

نحسب C_{Max} :

$$C_{Max}=\sqrt{\frac{L-1}{2}}=\sqrt{\frac{2-1}{2}}=\sqrt{0.5}=0.7$$

نلاحظ أن قيمة معامل التوافق C تبتعد عن C_{Max} وبالتالي العلاقة بين المتغيرين علاقة ضعيفة.

VIII–4. الارتباط بين متغيرين كميين :

VIII–4–1. معامل الارتباط :

قام العديد من العلماء بالبحث عن علاقة الارتباط بين متغيرين كميين ، فمنهم العالمين بيرسون وسيبرمان وكل منهما سمي معامله بإسمه ، لكن معامل بيرسون هو معامل الارتباط الأكثر إستخداما .

إن معامل بيرسون للإرتباط يعطى بالقاعدة التالية :

$$r_p=r_{xy}=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{(\sum(x_i-\bar{x})^2)(\sum(y_i-\bar{y})^2)}}=\frac{\sum x_iy_i\ -\ n\bar{x}\bar{y}}{\sqrt{(\sum x_i{}^2-n\bar{x}^2)(\sum y_i{}^2-n\bar{y}^2)}}$$

مثال رقم (8–06): لتكن البيانات الموالية والتي تمثل علامات 12 طالب في مقياسي الاحصاء والرياضيات .

الجدول رقم (8–08): توزيع 12 طالب حسب علامات مقياسي الإحصاء والرياضيات.

علامة الاحصاء	10	5	9	12	14	8	6	16	18	14	12	8
علامة الرياضيات	9	4	8	10	11	9	8	12	18	10	13	8

المطلوب : هل هناك علاقة إرتباط بين علامة الاحصاء وعلامة الرياضيات ؟

الحل : بما أن المتغيرين كميين فإننا سندرس علاقة الارتباط بمعامل الارتباط لبيرسون

الجدول رقم (8-08): إستخدام جدول التوزيع في حساب معامل الارتباط

	علامة الاحصاء x_i	علامة الرياضيات y_i	$(x_i-\bar{x})$	$(y_i-\bar{y})$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$	x_iy_i	x_i^2	y_i^2
	10	9	-1	-1	1	1	1	90	100	81
	5	4	-6	-6	36	36	36	20	25	16
	9	8	-2	-2	4	4	4	72	81	64
	12	10	1	0	0	1	0	120	144	100
	14	11	3	1	3	9	1	154	196	121
	8	9	-3	-1	3	9	1	72	64	81
	6	8	-5	-2	10	25	4	48	36	64
	16	12	5	2	10	25	4	192	256	144
	18	18	7	8	56	49	64	324	324	324
	14	10	3	0	0	9	0	140	196	100
	12	13	1	3	3	1	9	156	144	169
	8	8	-3	-2	6	9	4	64	64	64
$\sum$	132	120			132	178	128	1452	1630	1328

$$r_{xy}=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{(\sum(x_i-\bar{x})^2)(\sum(y_i-\bar{y})^2)}}=\frac{132}{\sqrt{(178)(128)}}=0.87$$

$$r_{xy}=\frac{\sum x_i\,y_i \;-\; n\bar{x}\bar{y}}{\sqrt{(\sum x_i{}^2-n\bar{x}^2)\,(\sum y_i{}^2-n\bar{y}^2)}}=\frac{1452-12(11)(10)}{\sqrt{(1630-12(11)^2)(1328-12(10)^2)}}$$

$$=\frac{132}{\sqrt{(178)(128)}}=0.87$$

خواص معامل الارتباط : يتميز معامل الارتباط لبيرسون بالخواص التالية :

- هو قيمة عددية تتراوح ما بين [-1 1] ؛
- إذا كان $r_{xy}=-1$ فالعلاقة بين المتغيرين عكسية (سالبة) وتامة ؛
- إذا كان $r_{xy}=1$ فالعلاقة بين المتغيرين طردية (موجبة) وتامة ؛
- إذا كان $r_{xy}=0$ فلا توجد علاقة بين المتغيرين ؛
- إذا كان $-1<r_{xy}<0$ فالعلاقة بين المتغيرين عكسية (سالبة) ؛
- إذا كان $0<r_{xy}<1$ فالعلاقة بين المتغيرين موجبة (طردية) ؛
- كلما إقتربت قيمة r_{xy} من (-1) أو (1) يدل ذلك على أن هناك علاقة قوية بين المتغيرين ؛

- و كلما إقتربت قيمة r_{xy} من (0) يدل ذلك على أن هناك علاقة ضعيفة بين المتغيرين ؛

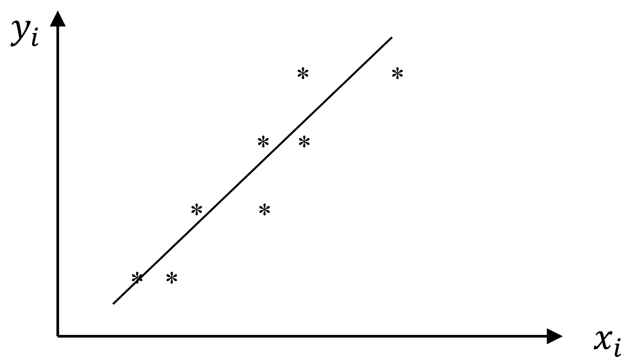
VIII-5. الانحدار:

لقد تطرقنا سابقا إلى المتغيرات الثنائية وطرق تحديد العلاقة بين متغيرين (كيفيين أو كميين)، ومن بين هذه الطرق معامل الارتباط r_p المستخدم في قياس وتحديد شكل العلاقة بين كتغيرين كميين ، فهو يقيس قوة أو ضعف العلاقة ، كما قد يحدد شكلها (عكسية أو طردية) ، فالعلاقة العكسية تعني أنه إذا إرتفع المتغير المستقل يؤدي إلى إنخفاض المتغير التابع والعكس ، مثل علاقة سعر سلعة ما (p_i) بكميتها المباعة (q_i) ، أما العلاقة الطردية (الموجبة) فتعني إذا إرتفع المتغير المستقل يؤدي إلى إرتفاع المتغير التابع والعكس ، مثل علاقة الدخل (y_i) بالاستهلاك (c_i).

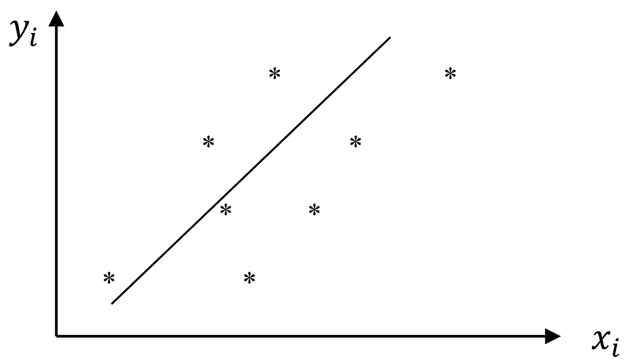
إن تحديد شكل العلاقة يمكن أن يتم بيانيا وهو ما يسمى بالانحدار ، أي إنحدار المتغير المستقل على المتغير التابع ، ويكون ذلك من خلال تحديد سحابة النقاط ثم تقدير معادلة الانحدار (الانحدار البسيط) .

VIII-5-1. سحابة النقاط : نعني بسحابة النقاط والتي تخص متغيرين كميين بإسقاط إحداثيات (تغيرات) المتغيرين التابع والمستقل على منحنى بياني ، حيث المحور الأفقي يخصص للمتغير المستقل (x_i) والمحور العمودي يخصص للمتغير التابع (y_i) ، فسحابة النقاط قد تأخذ عدة أشكال ، وكل شكل نستخلص منه شكل العلاقة وقوتها ، كما سنوضحه في المخططات الموالية :

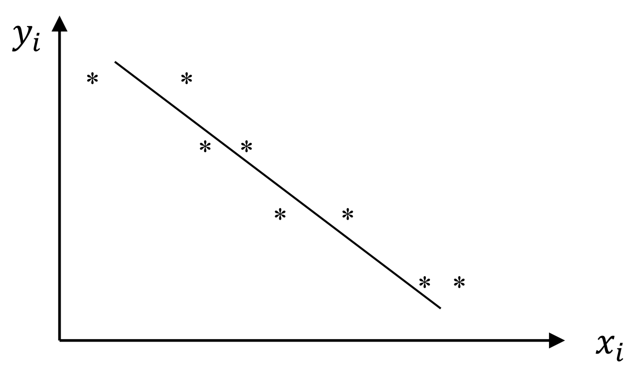
الشكل رقم (8-02): الأشكال الممكنة لسحابة النقاط



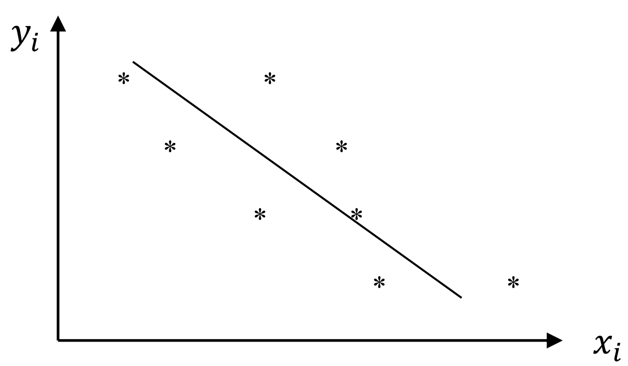
الشكل 01 : سحابة النقاط تقترب من خط المستقيم مما يشير إلى وجود علاقة طردية (موجبة) بين المتغيرين وقوية.



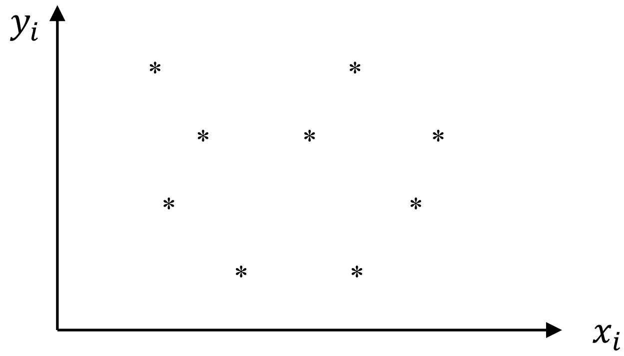
الشكل 02 : سحابة النقاط تبتعد من خط المستقيم مما يشير إلى وجود علاقة طردية (موجبة) بين المتغيرين وضعيفة.



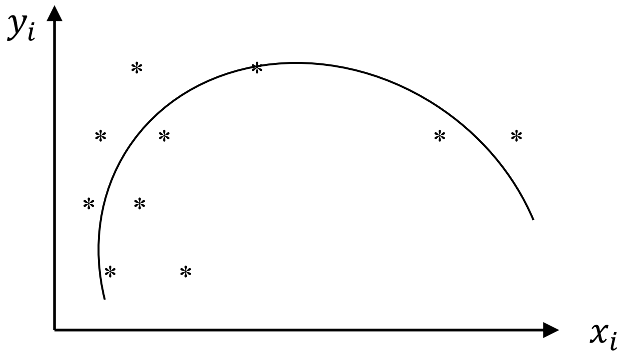
الشكل 03 : سحابة النقاط تقترب من خط المستقيم مما يشير إلى وجود علاقة عكسية (سالبة) بين المتغيرين وقوية.



الشكل 04 : سحابة النقاط تبتعد من خط المستقيم مما يشير إلى وجود علاقة عكسية (سالبة) بين المتغيرين وضعيفة.



الشكل 05 : سحابة النقاط لا يوجد لها إتجاه معين مما يشير إلى عدم وجود أي علاقة بين المتغيرين.

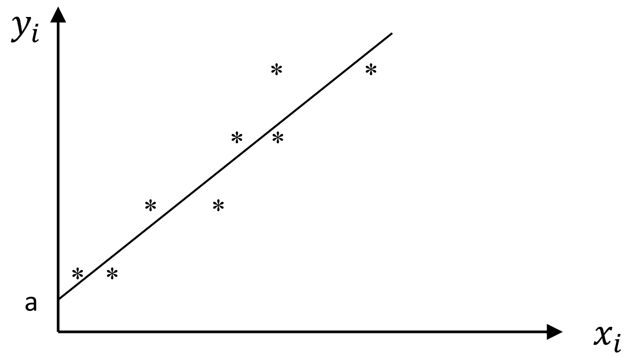


الشكل 06 : سحابة النقاط ليست في إتجاه واحد ،أي عدم وجود علاقة خطية بين المتغيرين .

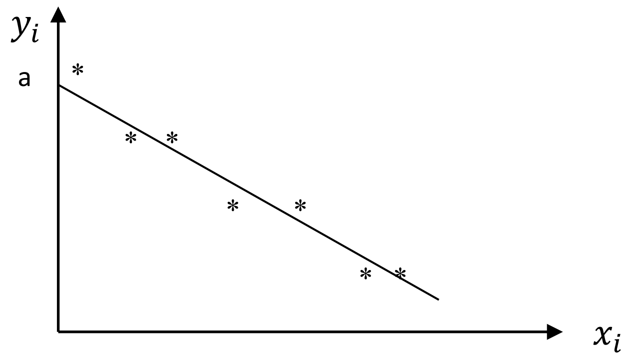
إذن من خلال الأشكال (سحابة النقاط) يمكن تحديد شكل العلاقة (خطية أو خطية ، طردية أو عكسية) ،كما يمكن تحديد قوة العلاقة أو ضعفها بين المتغيرين المستقل (x_i) و التابع (y_i).

VIII-5-2. الانحدار الخطي البسيط : يعرف الانحدار الخطي البسيط على أنه نموذج رياضي (أو علاقة رياضية) بين متغيرين كميين تتميز سحابة إنتشار العلاقة بينهما أنها عبارة عن خط أو مجال مستقيم وفق إحدى الشكلين :

الشكل رقم (8-03): الأشكال الممكنة للانحدار الخطي البسيط.



الشكل 01 : سحابة النقاط على شكل خط المستقيم مما يشير إلى وجود علاقة خطية وطردية (موجبة) بين المتغيرين.



الشكل 02 : سحابة النقاط على شكل خط المستقيم مما يشير إلى وجود علاقة خطية وعكسية (سالبة) بين المتغيرين.

بعد التأكد من شكل إنتشار النقاط الخطي بيانيا يتم تقدير العلاقة الخطية للنموذج الخطي البسيط وفق الصياغة الرياضية التالية :

$$y_i = a + bx_i$$

حيث : $b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$ (وهي تسمى بالميل)

و $a = \bar{y} - b\bar{x}$ (وهي نقطة تقاطع خط المستقيم مع المحور y_i)

ويستخدم النموذج المقدر عادة في التنبؤ بقيم المتغير التابع في المستقبل وهذا بعد معرفة القيمة المستقبلية للمتغير المستقل ، فمثلا إذا كانت بيانات المتغيرين من 2000 إلى 2022 ، فيمكن التنبؤ بقيمة المتغير التابع لسنة 2023 وهذا بعد معرفة قيمة تغير المتغير المستقل سنة 2023 ويتم تعويض ذلك في معادلة التقدير كما يلي :

$$y_{2023} = a + bx_{2023}$$

ملاحظة : يوجد العديد من نماذج الانحدار غير الخطي والتي من بينها القطع المكافئ أو المعادلة اللوغارتمية أو الأسية ، لكننا سنكتفي فقط بدراسة العلاقة الخطية بين المتغيرين.

مثال رقم (8-07): نعود إلى بيانات المثال السابق والخاص بإستهلاك إحدى العائلات ودخلها المتاح خلال الفترة 2010-2022 المرفقة في الجدول الموالي :

الجدول رقم (8-09): بيانات عن إستهلاك إحدى العائلات ودخلها المتاح الوحدة : 10^3 دينار

السنة	2010	2011	2012	2013	2014	2015	2016	2017	2018
الدخل المتاح	40	45	48	49	51	53	55	52	54
الاستهلاك	36	38	40	42	45	46	48	45	46

المطلوب :

- ما هو المتغير المستقل (المؤثر) والمتغير التابع (المتأثر) ؟
- ما هو شكل العلاقة بين المتغيرين ؟
- قدر أو أكتب شكل العلاقة بين المتغيرين رياضيا ؟
- ما هي قيمة إستهلاك هذه الأسرة لعام 2019 مع العلم أن دخلها المتاح سيرتفع بـ 10 % خلال نفس السنة ؟

الحل :

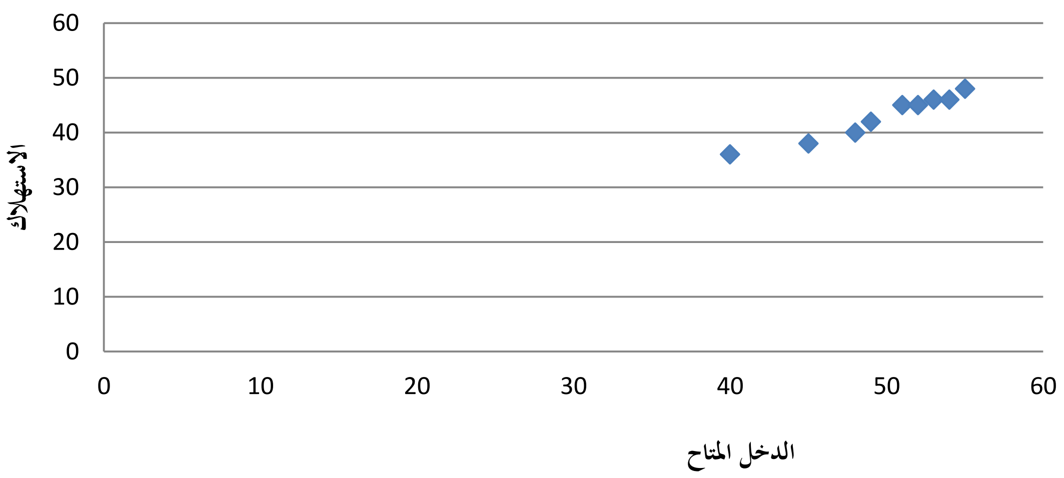
1- تقوم العائلة بقضاء حاجياتها (إستهلاكها) إنطلاقا من دخلها المتاح أي أن الدخل المتاح هو المتغير المؤثر (المستقل) والاستهلاك هو المتغير المتأثر (التابع) ، إذن :

✔ الدخل المتاح : X_i

✔ الاستهلاك : Y_i

2- تحديد شكل العلاقة : ويكون ذلك بيانيا كما يوضحه التمثيل الموالي

الشكل رقم (8-04): التمثيل البياني لشكل العلاقة بين الدخل المتاح والاستهلاك



نلاحظ أنه كلما زاد الدخل X_i زاد الاستهلاك Y_i عند هذه الأسرة ،كما أن سحابة النقاط تتجه على شكل مستقيم ، أي أن هناك علاقة خطية بين المتغيرين.

3- تقدير أو كتابة شكل العلاقة رياضيا :

$$y_i = a + bx_i$$

الجدول رقم (8-10): جدول التوزيع المستخدم في حساب العلاقة الانحدارية بين الدخل والاستهلاك

x_i^2	$x_i y_i$	الاستهلاك Y_i	الدخل X_i	السنة
1600	1440	36	40	2010
2025	1710	38	45	2011
2304	1920	40	48	2012
2401	2058	42	49	2013
2601	2295	45	51	2014
2809	2438	46	53	2015
3136	2688	48	56	2016
2809	2385	45	53	2017
3025	2585	47	55	2018
22710	19519	387	450	المجموع

$$\bar{x} = \frac{\sum x_i}{\sum n_i} = \frac{450}{9} = 50$$

$$\bar{y} = \frac{\sum y_i}{\sum n_i} = \frac{387}{9} = 43$$

$$b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{19519 - (9)(50)(43)}{22710 - 9(50^2)} = \frac{169}{210} = 0.8$$

$$a = \bar{y} - b\bar{x} = 43 - (0.8)(50) = 3$$

إذن : $y_i = 3 + 0.8x_i$

التنبؤ بإستهلاك الأسرة لعام 2019 ، أي y_{2019}

الدخل سيزيد عام 2019 بـ 10 % أي :

$$x_{2019} = x_{2018} + (0.1)x_{2018} = (1 + 0.1)x_{2018} = 1.1x_{2018} = 1.1(55) = 60.5$$

$$x_{2019} = 60.5$$

إذن :

$$y_{2019} = 3 + 0.8x_{2019} = 3 + 0.8(60.5) = 51.4$$

$$y_{2019} = 51.4$$

المراجع

المراجع

أولا : المراجع باللغة العربية

1- الموسى عبد الله عبد العزيز، استخدام الحاسوب الآلي في التعليم، مكتبة القشري، الرياض، 2001.

2- جلاطو جيلالي ، الإحصاء مع تمارين ومسائل محلولة ، ديوان المطبوعات الجامعية ، ط 05 ، 2005.

3- خالد أحمد فرحان المشهداني و رائد عبد الخالق عبد الله العبيدي، مبادئ الإحصاء : متضمن التحليل الاحصائي spss ، دار الأيام ، عمان، الأردن، 2013.

4- عبد الرزاق عزوز، الكامل في الإحصاء، الجزء الأول، ديوان المطبوعات الجامعية، الجزائر، 2010.

5- عبد الرزاق عزوز، الكامل في الإحصاء ، الجزء الثاني، ديوان المطبوعات الجامعية، الجزائر، 2011 .

6- عدنان عباس حميدان وآخرون ، مبادئ الإحصاء ، منشورات جامعة دمشق، سوريا ، 2016.

7- محمد بوهزة، محاضرات في الإحصاء الوصفي، دار المحمدية العامة، الجزائر، 2011.

8- محمد حسين محمد رشيد، الإحصاء : الوصفي والتطبيقي والحيوي ، دار صفاء ، عمان ، الأردن ، 2007.

9- محمد راتول، الإحصاء الوصفي، ديوان المطبوعات الجامعية، الجزائر، 2006.

ثانيا : المراجع باللغات الأجنبية

10 -BARBARA ILLOWSKY and SUSAN DEAN ; Introductory Statistics ; OpenStax ; Texas ; U.S.A ; 2013 .

11 - MURRAY R. SPIEGEL and LARRY J. STEPHENS ; Theory and Problems of STATISTICS ; Schaum's Outline Series ; Fourth Edition . 2008.